

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: April 17, 2005, 07:09:30 ; Search time 7160 Seconds  
(without alignments)  
5410.764 Million cell updates/sec

Title: US-09-671-687A-3  
Perfect score: 949  
Sequence: 1 MSSGLWSQEKVTSPLYEERI.....RLLCDAYNCWQSPMTSLYK 949

Scoring table: OLIGO  
Xgapop 60.0 , Xgapext 60.0  
Ygapop 60.0 , Ygapext 60.0  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 45554873 seqs, 20411521753 residues

Word size: 1  
Total number of hits satisfying chosen parameters: 91084298

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:  
-MODEL=frame+p2n.model -DEV=xlp  
-Q=/cgn2\_1/USPTO\_spool\_p/US09671687/runat\_15042005\_163039\_12012/app\_query.fasta\_1.1095  
-DB=Pending Patents NA Main -QFMT=fastap -SUFFIX=Oligo.rnpsm -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=oligo  
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=quality -THR MIN=1  
-ALIGN=15 -MODE=LOCAL -OUTFMT=ptco -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=2000000000 -USER=US09671687@cgn\_1\_1 6376 @runat\_15042005\_163039\_12012  
-NCPU=6 -ICPU=3 -NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100  
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : Pending Patents NA Main:  
1: /cgn2\_6/ptodata/1/pna/PCTUS1\_COMB.seq:  
2: /cgn2\_6/ptodata/1/pna/PCTUS2\_COMB.seq:  
3: /cgn2\_6/ptodata/1/pna/PCTUS\_COMB.seq:  
4: /cgn2\_6/ptodata/1/pna/US06\_COMB.seq:  
5: /cgn2\_6/ptodata/1/pna/US07\_COMB.seq:  
6: /cgn2\_6/ptodata/1/pna/US080\_COMB.seq:  
7: /cgn2\_6/ptodata/1/pna/US081\_COMB.seq:  
8: /cgn2\_6/ptodata/1/pna/US082\_COMB.seq:  
9: /cgn2\_6/ptodata/1/pna/US083\_COMB.seq:  
10: /cgn2\_6/ptodata/1/pna/US084\_COMB.seq:  
11: /cgn2\_6/ptodata/1/pna/US085\_COMB.seq:  
12: /cgn2\_6/ptodata/1/pna/US086\_COMB.seq:  
13: /cgn2\_6/ptodata/1/pna/US087\_COMB.seq:  
14: /cgn2\_6/ptodata/1/pna/US088\_COMB.seq:  
15: /cgn2\_6/ptodata/1/pna/US089\_COMB.seq:  
16: /cgn2\_6/ptodata/1/pna/US090\_COMB.seq:  
17: /cgn2\_6/ptodata/1/pna/US091\_COMB.seq:  
18: /cgn2\_6/ptodata/1/pna/US092A\_COMB.seq:  
19: /cgn2\_6/ptodata/1/pna/US092B\_COMB.seq:  
20: /cgn2\_6/ptodata/1/pna/US093A\_COMB.seq:  
21: /cgn2\_6/ptodata/1/pna/US093B\_COMB.seq:  
22: /cgn2\_6/ptodata/1/pna/US094\_COMB.seq:  
23: /cgn2\_6/ptodata/1/pna/US095A\_COMB.seq:  
24: /cgn2\_6/ptodata/1/pna/US095B\_COMB.seq:  
25: /cgn2\_6/ptodata/1/pna/US095C\_COMB.seq:  
26: /cgn2\_6/ptodata/1/pna/US095D\_COMB.seq:  
27: /cgn2\_6/ptodata/1/pna/US096A\_COMB.seq:  
28: /cgn2\_6/ptodata/1/pna/US096B\_COMB.seq:

29: /cgn2\_6/ptodata/1/pna/US096C\_COMB.seq:  
30: /cgn2\_6/ptodata/1/pna/US096D\_COMB.seq:  
31: /cgn2\_6/ptodata/1/pna/US096E\_COMB.seq:  
32: /cgn2\_6/ptodata/1/pna/US097A\_COMB.seq:  
33: /cgn2\_6/ptodata/1/pna/US097B\_COMB.seq:  
34: /cgn2\_6/ptodata/1/pna/US097C\_COMB.seq:  
35: /cgn2\_6/ptodata/1/pna/US098A\_COMB.seq:  
36: /cgn2\_6/ptodata/1/pna/US098B\_COMB.seq:  
37: /cgn2\_6/ptodata/1/pna/US098C\_COMB.seq:  
38: /cgn2\_6/ptodata/1/pna/US098D\_COMB.seq:  
39: /cgn2\_6/ptodata/1/pna/US099A\_COMB.seq:  
40: /cgn2\_6/ptodata/1/pna/US099B\_COMB.seq:  
41: /cgn2\_6/ptodata/1/pna/US099C\_COMB.seq:  
42: /cgn2\_6/ptodata/1/pna/US099D\_COMB.seq:  
43: /cgn2\_6/ptodata/1/pna/US099E\_COMB.seq:  
44: /cgn2\_6/ptodata/1/pna/US099F\_COMB.seq:  
45: /cgn2\_6/ptodata/1/pna/US099G\_COMB.seq:  
46: /cgn2\_6/ptodata/1/pna/US100A\_COMB.seq:  
47: /cgn2\_6/ptodata/1/pna/US100B\_COMB.seq:  
48: /cgn2\_6/ptodata/1/pna/US101A\_COMB.seq:  
49: /cgn2\_6/ptodata/1/pna/US101B\_COMB.seq:  
50: /cgn2\_6/ptodata/1/pna/US102A\_COMB.seq:  
51: /cgn2\_6/ptodata/1/pna/US102B\_COMB.seq:  
52: /cgn2\_6/ptodata/1/pna/US103A\_COMB.seq:  
53: /cgn2\_6/ptodata/1/pna/US103B\_COMB.seq:  
54: /cgn2\_6/ptodata/1/pna/US104A\_COMB.seq:  
55: /cgn2\_6/ptodata/1/pna/US104B\_COMB.seq:  
56: /cgn2\_6/ptodata/1/pna/US105A\_COMB.seq:  
57: /cgn2\_6/ptodata/1/pna/US105B\_COMB.seq:  
58: /cgn2\_6/ptodata/1/pna/US106A\_COMB.seq:  
59: /cgn2\_6/ptodata/1/pna/US107A\_COMB.seq:  
60: /cgn2\_6/ptodata/1/pna/US107B\_COMB.seq:  
61: /cgn2\_6/ptodata/1/pna/US107C\_COMB.seq:  
62: /cgn2\_6/ptodata/1/pna/US107D\_COMB.seq:  
63: /cgn2\_6/ptodata/1/pna/US108A\_COMB.seq:  
64: /cgn2\_6/ptodata/1/pna/US108B\_COMB.seq:  
65: /cgn2\_6/ptodata/1/pna/US109A\_COMB.seq:  
66: /cgn2\_6/ptodata/1/pna/US109B\_COMB.seq:  
67: /cgn2\_6/ptodata/1/pna/US109C\_COMB.seq:  
68: /cgn2\_6/ptodata/1/pna/US110\_COMB.seq:  
69: /cgn2\_6/ptodata/1/pna/US6000\_COMB.seq:  
70: /cgn2\_6/ptodata/1/pna/US6001\_COMB.seq:  
71: /cgn2\_6/ptodata/1/pna/US6002\_COMB.seq:  
72: /cgn2\_6/ptodata/1/pna/US6003\_COMB.seq:  
73: /cgn2\_6/ptodata/1/pna/US6004\_COMB.seq:  
74: /cgn2\_6/ptodata/1/pna/US6005\_COMB.seq:  
75: /cgn2\_6/ptodata/1/pna/US6006\_COMB.seq:  
76: /cgn2\_6/ptodata/1/pna/US6007\_COMB.seq:  
77: /cgn2\_6/ptodata/1/pna/US6008\_COMB.seq:  
78: /cgn2\_6/ptodata/1/pna/US6009\_COMB.seq:  
79: /cgn2\_6/ptodata/1/pna/US6010\_COMB.seq:  
80: /cgn2\_6/ptodata/1/pna/US6011\_COMB.seq:  
81: /cgn2\_6/ptodata/1/pna/US6012\_COMB.seq:  
82: /cgn2\_6/ptodata/1/pna/US6013\_COMB.seq:  
83: /cgn2\_6/ptodata/1/pna/US6014\_COMB.seq:  
84: /cgn2\_6/ptodata/1/pna/US6015\_COMB.seq:  
85: /cgn2\_6/ptodata/1/pna/US6016\_COMB.seq:  
86: /cgn2\_6/ptodata/1/pna/US6017\_COMB.seq:  
87: /cgn2\_6/ptodata/1/pna/US6018\_COMB.seq:  
88: /cgn2\_6/ptodata/1/pna/US6019\_COMB.seq:  
89: /cgn2\_6/ptodata/1/pna/US6020\_COMB.seq:  
90: /cgn2\_6/ptodata/1/pna/US6021\_COMB.seq:  
91: /cgn2\_6/ptodata/1/pna/US6022\_COMB.seq:  
92: /cgn2\_6/ptodata/1/pna/US6023A\_COMB.seq:  
93: /cgn2\_6/ptodata/1/pna/US6023B\_COMB.seq:  
94: /cgn2\_6/ptodata/1/pna/US6024\_COMB.seq:  
95: /cgn2\_6/ptodata/1/pna/US6025\_COMB.seq:  
96: /cgn2\_6/ptodata/1/pna/US6026\_COMB.seq:  
97: /cgn2\_6/ptodata/1/pna/US6027\_COMB.seq:  
98: /cgn2\_6/ptodata/1/pna/US6028\_COMB.seq:  
99: /cgn2\_6/ptodata/1/pna/US6029\_COMB.seq:  
100: /cgn2\_6/ptodata/1/pna/US6030\_COMB.seq:  
101: /cgn2\_6/ptodata/1/pna/US6031\_COMB.seq:

102: /cgn2\_6/ptodata/1/pna/US6032 COMB.seq.\*  
 103: /cgn2\_6/ptodata/1/pna/US6033 COMB.seq.\*  
 104: /cgn2\_6/ptodata/1/pna/US6034 COMB.seq.\*  
 105: /cgn2\_6/ptodata/1/pna/US6035 COMB.seq.\*  
 106: /cgn2\_6/ptodata/1/pna/US6036 COMB.seq.\*  
 107: /cgn2\_6/ptodata/1/pna/US6037 COMB.seq.\*  
 108: /cgn2\_6/ptodata/1/pna/US6038 COMB.seq.\*  
 109: /cgn2\_6/ptodata/1/pna/US6039 COMB.seq.\*  
 110: /cgn2\_6/ptodata/1/pna/US6040 COMB.seq.\*  
 111: /cgn2\_6/ptodata/1/pna/US6041 COMB.seq.\*  
 112: /cgn2\_6/ptodata/1/pna/US6042 COMB.seq.\*  
 113: /cgn2\_6/ptodata/1/pna/US6043 COMB.seq.\*  
 114: /cgn2\_6/ptodata/1/pna/US6044 COMB.seq.\*  
 115: /cgn2\_6/ptodata/1/pna/US6045 COMB.seq.\*  
 116: /cgn2\_6/ptodata/1/pna/US6046 COMB.seq.\*  
 117: /cgn2\_6/ptodata/1/pna/US6047 COMB.seq.\*  
 118: /cgn2\_6/ptodata/1/pna/US6048 COMB.seq.\*  
 119: /cgn2\_6/ptodata/1/pna/US6049 COMB.seq.\*  
 120: /cgn2\_6/ptodata/1/pna/US6050 COMB.seq.\*  
 121: /cgn2\_6/ptodata/1/pna/US6051 COMB.seq.\*  
 122: /cgn2\_6/ptodata/1/pna/US6052 COMB.seq.\*  
 123: /cgn2\_6/ptodata/1/pna/US6053 COMB.seq.\*  
 124: /cgn2\_6/ptodata/1/pna/US6054 COMB.seq.\*  
 125: /cgn2\_6/ptodata/1/pna/US6055 COMB.seq.\*  
 126: /cgn2\_6/ptodata/1/pna/US6056 COMB.seq.\*  
 127: /cgn2\_6/ptodata/1/pna/US6057 COMB.seq.\*  
 128: /cgn2\_6/ptodata/1/pna/US6058 COMB.seq.\*  
 129: /cgn2\_6/ptodata/1/pna/US6059 COMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	841	88.6	3715	US-09-671-687A-2	Sequence 2, Appli
2	754	79.5	4661	US-09-522-303-1293	Sequence 1293, Ap
3	754	79.5	4664	US-09-315-788-1729	Sequence 1729, Ap
4	754	79.5	4664	US-09-315-788A-1729	Sequence 1729, Ap
5	754	79.5	4664	US-09-783-514-1729	Sequence 1729, Ap
6	754	79.5	4664	US-09-783-514A-1729	Sequence 1729, Ap
7	754	79.5	4668	US-09-396-087-4309	Sequence 4309, Ap
8	754	79.5	4668	US-09-396-970-7559	Sequence 7559, Ap
9	754	79.5	4668	US-09-397-424-5019	Sequence 5019, Ap
10	754	79.5	4668	US-09-397-424A-5019	Sequence 5019, Ap
11	754	79.5	4668	US-09-432-241A-3760	Sequence 3760, Ap
12	754	79.5	4668	US-09-434-737-1278	Sequence 1278, Ap
13	754	79.5	4668	US-09-850-118-1278	Sequence 1278, Ap
14	754	79.5	6831	PCR-US02-27777-49	Sequence 49, Appl
15	754	79.5	6831	PCR-US02-27777A-49	Sequence 49, Appl
16	754	79.5	6871	US-10-170-235-39761	Sequence 39761, A
17	751	79.1	6314	US-09-324-185-23812	Sequence 23812, A
18	711	74.9	4527	US-09-786-797B-25	Sequence 25, Appl
19	711	74.9	4527	US-10-921-707-25	Sequence 25, Appl
20	711	74.9	4527	US-10-131-321-21	Sequence 21, Appl
21	685	72.2	5371	PCR-US02-14570-3	Sequence 3, Appli
22	685	72.2	5371	US-09-851-673-3	Sequence 3, Appli
23	685	72.2	5371	US-10-755-889-489	Sequence 489, App
24	685	72.2	5371	US-10-887-553A-199	Sequence 199, App
25	685	72.2	5371	US-10-440-068-489	Sequence 489, App
26	685	72.2	5371	US-10-469-757-489	Sequence 489, App
27	685	72.2	5371	US-10-172-360-22693	Sequence 22693, A
28	641	67.5	2523	US-09-488-725A-613	Sequence 613, App
29	641	67.5	2523	US-10-037-270-290	Sequence 290, App
30	641	67.5	2523	US-10-105-837-290	Sequence 290, App
31	641	67.5	2523	US-10-117-722-290	Sequence 290, App
32	641	67.5	2523	US-10-258-898A-613	Sequence 613, App
33	641	67.5	2523	US-10-286-898A-613	Sequence 613, App
34	641	67.5	4716	PCR-US01-01239-843	Sequence 843, App
35	641	67.5	4716	US-09-764-902-843	Sequence 843, App

## ALIGNMENTS

## RESULT 1

US-09-671-687A-2  
 ; Sequence 2, Application US/09671687A  
 ; GENERAL INFORMATION:  
 ; APPLICANT: WALLACH, David  
 ; APPLICANT: KOVALENKO, Andrei  
 ; APPLICANT: CANTARELLA, Giuseppe  
 ; TITLE OF INVENTION: INHIBITOR OF NF-KB ACTIVATOR  
 ; FILE REFERENCE: WALLACH-25  
 ; CURRENT APPLICATION NUMBER: US/09/671,687A  
 ; PRIOR FILING DATE: 2000-09-28  
 ; PRIOR APPLICATION NUMBER: 09/646,403  
 ; PRIOR FILING DATE: 2000-09-18  
 ; PRIOR APPLICATION NUMBER: IL 126024  
 ; PRIOR FILING DATE: 1998-09-01  
 ; PRIOR APPLICATION NUMBER: IL 134604  
 ; PRIOR FILING DATE: 2000-02-17  
 ; NUMBER OF SEQ ID NOS: 4  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 2  
 ; LENGTH: 3715  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: misc feature  
 ; LOCATION: (18)..(18)  
 ; OTHER INFORMATION: n is either a, c, g, or t.  
 ; NAME/KEY: misc feature  
 ; LOCATION: (22)..(22)  
 ; OTHER INFORMATION: n is either a, c, g, or t.  
 ; NAME/KEY: misc feature  
 ; LOCATION: (756)..(756)  
 ; OTHER INFORMATION: n is either a, c, g, or t.  
 ; NAME/KEY: misc feature  
 ; LOCATION: (1348)..(1348)  
 ; OTHER INFORMATION: n is either a, c, g, or t.  
 ; US-09-671-687A-2

Alignment Scores:  
 Pred. No.: 0 Length: 3715  
 Score: 841.00 Matches: 949  
 Percent Similarity: 98.7% Conservative: 0  
 Best Local Similarity: 98.7% Mismatches: 0  
 Query Match: 88.6% Indels: 12  
 DB: 30 Gaps: 0

US-09-671-687A-3 (1-949) x US-09-671-687A-2 (1-3715)

QY 1 MetSerSerGlyLeuTrpSerGlnGlnLysValThrSerProTyrTrpGluAlaArgIle 20  
 Db 497 ATGAGTTCAGGCTTATGGAGCCAGCAAAAAGTCACTTCACCTACTGGGAGAGCGGATT 556  
 QY 21 PheTyrLeuLeuLeuGlnGlnCysSerValThrAspLysGlnThrGlnLysLeuLeuLys 40  
 Db 557 TTTTACTTGTCTTCITCAAGAATGTCAGCGTTACAGCAAAACAAACAAAGCTCTCTAAA 616  
 QY 41 ValProLysGlySerIleGlyGlnTrpIleGlnAspArgSerValGlyHisSerArgIle 60  
 Db 617 GTACCGAGGAGNAGTATAGGACAGTATATTCAAGATCGTTCTGTGGGCGATTCAAGGATT 676

Sequence 5981, Ap  
 Sequence 7384, Ap  
 Sequence 18842, A  
 Sequence 18842, A  
 Sequence 1878, Ap  
 Sequence 4528, Ap  
 Sequence 1878, Ap  
 Sequence 4528, Ap  
 Sequence 5009, Ap  
 Sequence 5009, Ap

QY 61 ProSerAlaLysGlyLysAsnGlnIleGlyLeuLysIleLeuGluGlnProHisAla 80  
DB 677 CTTCTGCAAAAGGCAAGAAATCAGATTGGATTAAAAATTTCTAGAGCAACTCATGCA 736  
QY 81 ValLeuPheValAspGlu---AspValValGluIleAsnGluLysPheThrGluLeuLeu 99  
DB 737 GTTCTCTTTGATGAAANGATGTTGTAGAGATAAATGAAAGTTTCACAGAGTTACTT 796  
QY 100 LeuAlaIleThrAsnCysGluGluArgPheSerLeuPheLysAsnArgAsnArgLeuSer 119  
DB 797 TTGGCAATTACCAATTTGAGAGAGGTTACGCTGTTTAAAAACAGAAACAGACTAAGT 856  
QY 120 LysGlyLeuGlnIleAspValGlyCysProValLysValGlnLeuArgSerGlyGluGlu 139  
DB 857 AAAGCCTCCAAATAGACGTGGCTGCTCTGTGAAAGTACACTGAGATCTGGGGAAGAA 916  
QY 140 LysPheProGlyValValArgPheArgGlyProLeuLeuAlaGluArgThrValSerGly 159  
DB 917 AAATTTCTGAGTTGTACGCTTCAGAGGACCCCTGTTAGCAGAGAGACAGTCTCCGGA 976  
QY 160 IlePhePheGlyValGluLeuLeuGluGlyArgGlyGlnGlyPheThrAspGlyVal 179  
DB 977 ATATTCTTTGAGTTGAAATTTGCTGGAAGAGTCTGTTCAAGGTTTCACTGACGGGTG 1036  
QY 180 TyrGlnGlyLysGlnLeuPheGlnCysAspGluAspCysGly---PheValAlaLeuAsp 198  
DB 1037 TACCAGGGAACAGCTTTTTCAGTGTGATGAAGATTGGCGTGTGTTGTCATTGGAC 1096  
QY 199 LysLeuGluLeuIleGluAspAspThrAlaLeuGluSerAspTyrAlaGlyProGly 218  
DB 1097 AAGCTAGAACTCATAGAGATGATGACACTGCATTGGAAAGTATTACGACAGTCTCTGG 1156  
QY 219 AspThrMetGlnValGluLeuProLeuGluIleAsnSerArgValSerLeuLysGly 238  
DB 1157 GACACAAATGCAAGTCTGAACTTCTCTCTTTGAAATAAACTCCAGAGTTTCTTTGAAGGT 1216  
QY 239 GlyGluThrIleGluSerGlyThrValIlePheCysAspValLeuProGlyLysGluSer 258  
DB 1217 GGAGAAACAATAGAACTGGAAACAGTTATATCTGTGATGTTTCCAGGAAAGAAAGC 1276  
QY 259 LeuGlyTyrPheValGlyValAspMetAspAsnProIleGlyAsnTyrAspGlyArgPhe 278  
DB 1277 TTAGGATATTTGTTGGTGTGACATGATGATTAACCTATTGGCACTGGATGGAAGATT 1336  
QY 279 AspGlyVal---LeuCysSerPheAlaCysValGluSerThrIleLeuLeuHisIleAsn 297  
DB 1337 GATGGAGTGCANCTTTGTAGTTTTCGCTGTGTTGAAAGTACAAATCTATTGACATCAAT 1396  
QY 298 AsplleleProGluSerValThrGlnGluArgArgProProLysLeuAlaPheMetSer 317  
DB 1397 GATATCATCCAGAGAGTGTGACGAGAAAGGAGGCTCCCAAACTTGCCTTTATATGTCA 1456  
QY 318 ArgGlyValGlyAspLysGlySerSerHisLeuLysProLysAlaThrGlySerThr 337  
DB 1457 AGAGTGTGGGCAAGGTTTCCAGTCAATAATAAACCAAGGCTACAGGATCTACC 1516  
QY 338 SerAspProGlyAsnArg---ArgSerGluLeuPheTyrThrLeuAsnGlySerSerVal 356  
DB 1517 TCAGACCTCGAATAGAAACAGATCTCAATATATTTATACCTTAATGGGCTTCTGTT 1576  
QY 357 AspSerGlnProGlnSerLysSerLysAsnThrTyrTyrIleAspGluValAlaGluAsp 376  
DB 1577 GACTCAACACCAATCCAAATCAAAAAATACATGTTGATGAAAGTTCAGAGAGAC 1636  
QY 377 ProAlaLysSerLeuThrGluIleSerThrAspPheAspArgSerSerProLeuGln 396  
DB 1637 CCTGCAAAATCTTTACAGAGATATACAGACTTTGACCGTCTTCCACCACCTCCAG 1696  
QY 397 ProProValAsnSerLeuThrThrGluAsnArgPheHisSerLeuProPheSerLeu 416  
DB 1697 CCTCTCTCTGTAACCTCACTGACCCAGCAACAGATTCCACTCTTTACCATTCAGTCTC 1756

QY 417 ThrLysMetProAsnThrAsnGlySerIleGlyHisSerProLeuSerLeuSerAlaGln 436  
DB 1757 ACCAAGATGCCCAATACCAATGGAAGTATTGGCCACAGTCCACTTTCTCTGTGAGCCAG 1816  
QY 437 SerValMetGluGluLeuAsnThrAlaProValGlnGluSerProProLeuAlaMetPro 456  
DB 1817 TCTGTAATGGAAGAGCTAAACACTGCACCGCTGCAAGAGAGTCCACCTTTGGCCATGCCT 1876  
QY 457 ProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLysGluAsnProPro 476  
DB 1877 CTGGGAATCTACATGCTTAGAAGTGGGCTCAATGGCTGAAAGTAAAGAGAACCTCTCT 1936  
QY 477 PheTyrGlyValIleArgTyrIleGlyGlnProProGlyLeuAsnGluValLeuAlaGly 496  
DB 1937 TTTCTATGGGTAAATCCGTTGGATCGGTAGCCACACAGAGTGAATGAAGTCTCGCTGGA 1996  
QY 497 LeuGluLeuGluAspGluCysAlaGlyCysThrAspGlyThrPheArgGlyThrArgTyr 516  
DB 1997 CTGGAACCTGAAGATGAGTGTGAGGCTGTACGATGGAACCTTCAGAGCACTCGGTAT 2056  
QY 517 PheThrCysAlaLeuLysLysAlaLeuPheValLysLeuLysSerCysArgProAspSer 536  
DB 2057 TTTCCCTGTGCCCTGGAAGGCGCTGTTTGTGAAACTGAAGAGCTGCAAGGCTGACTCT 2116  
QY 537 ArgPheAlaSerLeuGlnProValSerAsnGlnIleGluArgCysAsnSerLeuAlaPhe 556  
DB 2117 AGTTTTCATCATTTGAGCCGCTTCCAAATCAGATTGAGCGCTGTAACTCTTTAGCATTT 2176  
QY 557 GlyGlyTyrLeuSerGluValValGluGluAsnThrProProLysMetGluLysGluGly 576  
DB 2177 GGAGGCTACTTAAGTGAAGTAGTAGAAGAAATACTCCACCAAAATGGAAGAAAGGC 2236  
QY 577 LeuGluIleMetIleGlyLysLysGlyIleGlnGlyHisTyrAsnSerCysTyrLeu 596  
DB 2237 TTGGAGATAATGATTGGGAAGAAAGGATCATCGGTCATTACAAATCTCTGTTACTTA 2296  
QY 597 AspSerThrLeuPheCysLeuPheAlaPheSerSerValLeuAspThrValLeuLeuArg 616  
DB 2297 GACTCAACCTTATCTGCTTATTTGCTTTTAGTTCTGTTCTGACACTGTGTTACTTAGA 2356  
QY 617 ProLysGluLysAsnAspValGluTyrTyrSerGluThrGlnGluLeuArgThrGlu 636  
DB 2357 CCCAAAGAAAGAAACGATGATGATATATAGTGAACCCAGAGACTACTGAGAGACAGAA 2416  
QY 637 IleValAsnProLeuArgIleTyrGlyTyrValCysAlaThrLysIleMetLysLeuArg 656  
DB 2417 ATTGTTAATCTCTCAGAAATATGATATGTGTGTGCCACAAATAATTAAGAACTGAGG 2476  
QY 657 LysIleLeuGluLysValGluAlaAlaSerGlyPheThrSerGluGluLysAspProGlu 676  
DB 2477 AAAATACTTTGAAAGGTGGAGGCTGCATCAGGATTTTACCTCTGAAGAAAAGATCCTGAG 2536  
QY 677 GluPheLeuAsnIleLeuPheHisIleLeuArgValGluProLeuLeuLysIleArg 696  
DB 2537 GAATTTCTGAATATTCTGTTTTCATCATATTTAAGGTTAGAACCTTTGCTTAAATAAAGA 2596  
QY 697 SerAlaGlyGlnLysValGlnAspCysTyrPheTyrGlnIlePheMetGluLysAsnGlu 716  
DB 2597 TCAGCAGGTCAAAGGTACAGATTGTTTACTTCTATCAAAATTTTATGAAAAAATGAG 2656  
QY 717 LysValGlyValProThrIleGlnLeuLeuGluTyrPheSerPheIleAsnSerAsnLeu 736  
DB 2657 AAAAGTGGCGTTCCACAAATTCAGCAGTTGTTAGAAATGCTCTTTTATCAACAGTAACTCT 2716  
QY 737 LysPheAlaGluAlaProSerCysLeuIleIleGlnMetProArgPheGlyLysAspPhe 756  
DB 2717 AAATTTGAGAGGACCATCATGTTGATATTTCAGATGCTTCGATTTGGAAGAACTTTT 2776  
QY 757 LysLeuPheLysIlePhePheProSerLeuGluLeuAsnIleThrAspLeuLeuGluAsp 776  
DB 2777 AAACATTTTAAAAAATTTTCTCTCTGGAATTAATATAACAGATTTTACTTTGAAGAC 2836  
QY 777 ThrProArgGlnCysArgIleCysGlyLeuAlaMetTyrGluCysArgGluCysTyr 796

```
Db 2837 ACTCCACAGAGTCCGGATATGTGGAGGCTTGCAATGTATAGTGTAGAGATGCTAC 2896
Qy 797 AspAspProAapIleSerAlaGlyIleLysGlnPheCysLysThrCysAsnThrGln 816
Db 2897 GACCATCGGACATCTCAGCTGGAAAAATCAAGCAGTTTGTAAACCTGCCAACACTCAA 2956
Qy 817 ValHisLeuHisProLysArgLeuAsnHisLysTyrAsnProValSerLeuProLysAsp 836
Db 2957 GTCCACCTTCATCCGAGAGGCTGATCATATAATATACCCNGTGTCACTTCCCAAGAC 3016
Qy 837 LeuProAspTyrAspTyrArgHisGlyCysIleProCysGlnAsnMetGluLeuPheAla 856
Db 3017 TTACCCGACTGGGACTGGAGACACGGCTGCATCCCTTGCAGATATGGAGTTATTGCT 3076
Qy 857 ValLeuCysIleGluThrSerHisTyrValAlaPheValLysTyrGlyValAspAspSer 876
Db 3077 GTTCTCTGCATAGAAACAAGCCACTATGTGCTTTGTGAAGTATGGGAAGGACGATTTCT 3136
Qy 877 AlaTyrLeuPhePheAspSerMetAlaAspArgAspGlyGlyGlnAsnGlyPheAsnIle 896
Db 3137 GCCTGGCTCTTCTTGACAGCATGGCCGATCGGATGGTGTGTCAGAAATGGCTTCAACATT 3196
Qy 897 ProGlnValThrProCysProGluValGlyGluTyrLeuLysMetSerLeuGluAspLeu 916
Db 3197 CCTCAAGTCACCCCATGCCAGAGCTAGGAGAGTACTTGAAGATGCTCTCGAAGACCTG 3256
Qy 917 HisSerLeuAspSerArgArgIleGlnGlyCysAlaArgArgLeuLeuCysAspAlaTyr 936
Db 3257 CATTCCTTGGACTCCAGAGAAATCCAAAGCTGTGCACGAAGACTGCTTGTGATGCATAT 3316
Qy 937 MetCysMetTyrGlnSerProThrMetSerLeuTyrLys 949
Db 3317 ATGTGATGTACAGAGTCCAAACATGATGATTTGTACAAA 3355
```

## RESULT 2

```
US-09-522-303-1293
; Sequence 1293, Application US/09522303
; GENERAL INFORMATION:
; APPLICANT: Gearing, David P.
; APPLICANT: Holtzman, Douglas A.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES DERIVED FROM A
; TITLE OF INVENTION: HUMAN MAMMARY EPITHELIUM LIBRARY
; FILE REFERENCE: 1600.1086-001
; CURRENT APPLICATION NUMBER: US/09/522,303
; CURRENT FILING DATE: 2000-03-08
; EARLIER APPLICATION NUMBER: 60/123,393
; EARLIER FILING DATE: 1999-03-08
; NUMBER OF SEQ ID NOS: 1353
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1293
; LENGTH: 4661
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(4661)
; OTHER INFORMATION: n = A,T,C or G
US-09-522-303-1293
```

## Alignment Scores:

```
Pred. No.: 0 Length: 4661
Score: 754.00 Matches: 948
Percent Similarity: 98.54% Conservative: 0
Best Local Similarity: 98.54% Mismatches: 1
Query Match: 79.45% Indels: 14
DB: 23 Gaps: 0
```

US-09-671-687A-3 (1-949) x US-09-522-303-1293 (1-4661)

```
Qy 1 MetSerSerGlyLeuTyrSerGlnGlyLysValThrSerProTyrTyrGluGluArgIle 20
Db 231 ATGAGTTTCAGGCTTATGGAGCAAGAAAAGTCACTTCACCCCTACTGGGAGAGCGGATT 290
```

```
Qy 21 PheTyrLeuLeuLeuGlnGluCysSerValThrAspLysGlnThrGlnLysLeuLeuLys 40
Db 291 TTTTACTTGTCTCTTCAAGATGCGAGCTTACAGACAAACAAACAAACAAAGCTCTTAAA 350
Qy 41 ValProLysGlySerIleGlyGlnTyrIleGlnAspArgSerValGlyHisSerArgIle 60
Db 351 GTACCGAAGGGAAGTATAGGACAGTATATTCAAGATCGTCTCTGTGGGCAATTCAGGATT 410
Qy 61 ProSerAlaLysGlyLysLysAsnGlnIleGlyLeuLysIleLeuGluGlnProHisAla 80
Db 411 CCTCTGTCAGAAAGGCAAGAAAATCAGATTCGATTAAAAATTTCTAGACCACTCATGCA 470
Qy 81 ValLeuPheValAspGlu --- AspValValGluIleAsnGluLysPheThrGluLeuLeu 99
Db 471 GTTCTCTTTGTGTGAAAGAGATGTTGTAGATGAATGAATAAGTTTACAGAGTTACTT 530
Qy 100 LeuAlaIleThrAsnCysGluGluArgPheSerLeuPheLysAsnArgAsnArgLeuSer 119
Db 531 TTGGCAATTACCAATTGTGAGGAGAGGTTTACGCTGTTTAAAAACAGAAAACAGACTAAGT 590
Qy 120 LysGlyLeuGlnIleAspValGlyCysProValLysValGlnLeuArgSerGlyGluGlu 139
Db 591 AAAGCCCTCCAAATAGACGTGGCTGTCTGTGAAAGTACAGCTGAGATCTGGGGAAGAA 650
Qy 140 LysPheProGlyValValArgPheArgGlyProLeuLeuAlaGluArgThrValSerGly 159
Db 651 AAATTTCTGGAGTTGTACGCTTACAGAGACCCCTGTAGCAGAGAGACAGTCTCCGGA 710
Qy 160 IlePhePheGlyValGluLeuLeuGluGlyArgGlyGlnGlyPheThrAspGlyVal 179
Db 711 ATATTCTTTGGAGTTGAATTGCTGGAAGAGGTGCTGTGTCAGGTTTTCACCTGACGGGTG 770
Qy 180 TyrGlnGlyLysGlnLeuPheGlnCysAspGluAspCysGly --- PheValAlaLeuAsp 198
Db 771 TACCAAGGGAACACAGCTTTTTCAGTGTGATGAAGATTGTGGCGTGTGTTGTCATTGGAC 830
Qy 199 LysLeuLeuLeuLeuLeuAspAspThrAlaLeuGluSerAspTyrAlaGlyProGly 218
Db 831 AAGCTAGAACCTCATGAAGATGACACTGCTGGAAGAGTATGACAGGTCTCGGG 890
Qy 219 AspThrMetGlnValGluLeuProLeuGluIleAsnSerArgValSerLeuLysGly 238
Db 891 GACACAATGCGAGTGCNACTTCTCTTTGGAAATAAACTCCAGAGTTTCTTTGAA-GGT 949
Qy 239 -GlyGluThrIleGluSerGlyThrValIlePheCysAspValLeuProGlyLysGluSe 258
Db 950 TCGAGAAAACAATAGAAATCTGGAAACAGTTATATCTGTGATGTTTTCAGGAAAAGAAAG 1009
Qy 258 IleuGlyTyrPheValGlyValAspMetAspAsnProIleGlyAsnTyrAspGlyArgph 278
Db 1010 CTTAGGATATTTTGTGTGTGGACATGGAATACCTTATGGCAACTGGGATGGAAGATT 1069
Qy 278 eAspGlyVal --- LeuCysSerPheAlaCysValGluSerThrIleLeuLeuHisIleAs 297
Db 1070 TGATGAGTGCAGCTTTGTAGTTTTCGTGTGTGGAAGTACAAATCTATTGCAATCAA 1129
Qy 297 nAspIleIleProGluSerValThrGlnGluArgProProLysLeuAlaPheMetSe 317
Db 1130 TGATATCATCCAGAGAGTGTGACGACGAAAGAGGCGCTCCAAACTTGCCTTTATGTC 1189
Qy 317 rArgGlyValGlyAspLysGlySerSerSerHisAsnLysProLysAlaThrGlySerTh 337
Db 1190 AAGAGTGTGTGGGCAAAAGGTTTCATCCAGTCAATAATAACCAAAAGGCTTACAGGATCTAC 1249
Qy 337 rSerAspProGlyAsnArg --- ArgSerGluLeuPheTyrThrLeuAsnGlySerSerVa 356
Db 1250 CTCAGACCTTGGAAATAGAAACAGATCTGAATATTTTATACCTTAATGGGTCTTCTGT 1309
Qy 356 lAspSerGlnProGlnSerLysSerLysAsnThrTyrTyrIleAspGluValAlaGluAs 376
Db 1310 TGACTCACACACCACAATCCAAATCAAAAATAACATGTTGATGATGATGATGATGATGATGATG 1369
```



```
Qy 376 pProAlaLysSerLeuThrGluIleSerThrAspPheAspArgSerSerProProLeuG1 396
Db 1370 CCCTGC AAAATCTCTTACAGAGATATACAGACTTTGACCGTCTCTTACACCACTCCA 1429
Qy 396 nProProValAsnSerLeuThrGluAsnArgPheHisSerLeuProPheSerLe 416
Db 1430 GCCTCCTCTGTGAACCTCAGTACACCGAGAACAGATTCCACTCTTTACCACTTCACTCT 1489
Qy 416 uThrLysMetProAsnThrAsnGlySerIleGlyHisSerProLeuSerLeuSerAlaG1 436
Db 1490 CACCAAGATGCCAATACCAATGGAAGATTGGCCACAGCTCCACTTCTCTGTGAGCCCA 1549
Qy 436 nSerValMetGluGluLeuAsnThrAlaProValGlnGluSerProProLeuAlaMetPr 456
Db 1550 GTCTGTAATGAAGAGCTAAACACTGCACCCGCTCCAAAGAGATCCACCCCTGGCCATGCC 1609
Qy 456 oProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLysGluAsnProPr 476
Db 1610 TCCTGGGAACCTCACATGGTCTAGAAAGTGGGCTCATTTGGCTGAAGTTAAGGAGAACCCCTCC 1669
Qy 476 oPheTyrglyValIleArgTropIleGlyGlnProProGlyLeuAsnGluValLeuAlaG1 496
Db 1670 TTTCTATGGGGTAATCCGTGTGATCGGTACGACAGCTGAATGAAGTGTCTGCTGG 1729
Qy 496 yLeuGluLeuGluAspGluCysAlaGlyCysThrAspGlyThrPheArgGlyThrArgTy 516
Db 1730 ACTGGAACCTGGAAGATGAGTGTGACAGGCTGTACGGATGGAACCTTCAGAGGCACTCGGTA 1789
Qy 516 rPheThrCysAlaLeuLysLysAlaLeuPheValLysLeuLysSerCysArgProAspSe 536
Db 1790 TTTCACTGTGCTTGAAGAGGCGCTGTTGTGAACCTGAAGAGCTGCAGGCGCTGACTC 1849
Qy 536 rArgPheAlaSerLeuGlnProValSerAsnGlnIleGluArgCysAsnSerLeuAlaPh 556
Db 1850 TAGGTTTGCACTATTCGACCGCGTTTCCAATCAGATTGAGCGCTGTAACCTCTTTAGCA 1909
Qy 556 eGlyGlyTyrrLeuSerGluValValGluAsnThrProProLysMetGluLysGluG1 576
Db 1910 TGGAGGCTACTTAAGTGAAGTAGTAGAAGAAATACTCCACCAAAAATGGAAGAAAGG 1969
Qy 576 yLeuGluIleMetIleGlyLysLysGlyIleGlnGlyHisTyrrAsnSerCysTyrrLe 596
Db 1970 CTTGAGATAATGATTGGGAAGAAAGGATCCAGGCTCATTAACAATTTCTTTTACTT 2029
Qy 596 uAspSerThrLeuPheCysLeuPheAlaPheSerSerValLeuAspThrValLeuLeuAr 616
Db 2030 AGACTCAACCTTAATCTGCTTATTTGCTTTAGTCTTCTGACACTGTGTACTTAG 2089
Qy 616 gProLysGluLysAsnAspValGluTyrrTyrrSerGluThrGlnGluLeuLeuArgThrG1 636
Db 2090 ACCCAAGAAAGAACGATGAGATAATATATAGTGAACCCCAAGAGCTACTGAGGACAGA 2149
Qy 636 uIleValAsnProLeuArgIleTyrrGlyTyrrValCysAlaThrLysIleMetLysLeuAr 656
Db 2150 AATTGTTAATCCTCTGAGAATATATGATATGTGTGTCACCAAAAATTAAGAACTGAG 2209
Qy 656 gLysIleLeuGluLysValGluAlaSerGlyPheThrSerGluGluLysAspProG1 676
Db 2210 GAAAATCTTGAAGAGGTGGAGGCTGCATCAGGATTTACTCTGAAGAAAAGATCCTCA 2269
Qy 676 uGluPheLeuAsnIleLeuPheHisIleLeuArgValGluProLeuLeuLysIleAr 696
Db 2270 GGAATCTTGAATATTTCTGTTTCATCATATTTTAAGGGGTAGAACCTTTGCTAAAAATAAG 2329
Qy 696 gSerAlaGlyGlnLysValGlnAspCysTyrrPheTyrrGlnIlePheMetGluLysAsnG1 716
Db 2330 ATCAGCAGCTCAAAAGGTACAAGATTGTACTTCTATCAAAATTTTATGAAAAAATAAG 2389
Qy 716 uLysValGlyValProThrIleGlnLeuLeuGluTrpSerPheIleAsnSerAsnLe 736
Db 2390 GAAAGTTGGCGTTCCCAAAATTCAGCAGTGTGTAGAAATGGTCTTTTATCAACAGTAACCT 2449
Qy 736 uLysPheAlaGluAlaProSerCysLeuIleIleGlnMetProArgPheGlyLysAspPh 756
```

```
Db 2450 GAAATTTGCAGAGGCCACCATCATGTCTGATTATTTCAGATGCCTTCGATTTGGAAAAACCTT 2509
Qy 756 eLysLeuPheLysLysIlePheProSerLeuGluLeuAsnIleThrAspLeuLeuGluAs 776
Db 2510 TAAACATATTTAAAAAATTTTCTCTCTGGAATTAATATAACAGATTTACTTGAAGA 2569
Qy 776 pThrProArgGlnCysArgIleCysGlyGlyLeuAlaMetTyrrGluCysArgGluCysTy 796
Db 2570 CACTCCACAGACAGTGCAGGATATGTGAGGGCTTGCATATGATGATGATAGTAGAATGCTA 2629
Qy 796 rAspAspProAspIleSerAlaGlyLysIleLysGlnPheCysLysThrCysAsnThrG1 816
Db 2630 CGACGATCCGGACATCTCAGCTGGAATAATCAAGCAGTTTGTAAAAACCTGCACACTCA 2689
Qy 816 nValHisLeuHisProLysArgLeuAsnHisLysTyrrAsnProValSerLeuProLysAs 836
Db 2690 AGTCCACCTTTCATCCGAGAGGCTGAATCATATAATAACCCAGTGTCACTTCCCAAGA 2749
Qy 836 pLeuProAspTrpAspTrpArgHisGlyCysIleProCysGlnAsnMetGluLeuPheAl 856
Db 2750 CTTACCCGACTGGGACTGGAGACACGGCTGCATCCCTTGCACAGATATGGAGTTATTGC 2809
Qy 856 aValLeuCysIleGluThrSerHisTyrrValAlaPheValLysTyrrGlyLysAspSe 876
Db 2810 TGTCTCTCATAGAAAACAAGCCACTATGTGTCTTTGTGAAGTATGGGAAGGACGATTC 2869
Qy 876 rAlaTrpLeuPhePheAspSerMetAlaAspArgAspGlyGlyGlnAsnGlyPheAsnI1 896
Db 2870 TGCTTGGCTCTCTTTTGACAGCATGGCCGATCGGATGGTGGTCAAGATGGCTTCAACAT 2929
Qy 896 eProGlnValThrProCysProGluValGlyGluTyrrLeuLysMetSerLeuGluAspLe 916
Db 2930 TCTCTAAGTCAACCCATGCCAGAGTAGAGTAGTAGTCTTGAAGATGTCTCTGGAAGACCT 2989
Qy 916 uHisSerLeuAspSerArgArgIleGlnGlyCysAlaArgArgLeuLeuCysAspAlafy 936
Db 2990 GCATCTCTTGGACTCCAGGAGAAATCCAGGCTGTGCACGAAGACTGCTTTGTGATGCATA 3049
Qy 936 rMetCysMetTyrrGlnSerProThrMetSerLeuTyrrLys 949
Db 3050 TATGTGCATGTACCAGAGTCCAAACATGATGATTTGTACAAA 3089
```

## RESULT 3

```
US-09-315-788-1729
; Sequence 1729, Application US/09315788
; GENERAL INFORMATION:
; APPLICANT: Gearing, David P.
; APPLICANT: Holtzman, Douglas A.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES DERIVED FROM A
; TITLE OF INVENTION: HUMAN ASTROCYTE LIBRARY
; FILE REFERENCE: MLN98-19PM
; CURRENT APPLICATION NUMBER: US/09/315,788
; CURRENT FILING DATE: 1999-05-21
; EARLIER APPLICATION NUMBER: 60/086,455
; EARLIER FILING DATE: 1998-05-22
; EARLIER APPLICATION NUMBER: 60/132,067
; EARLIER FILING DATE: 1999-04-30
; NUMBER OF SEQ ID NOS: 2346
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1729
; LENGTH: 4664
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(4664)
; OTHER INFORMATION: n = A,T,C or G
US-09-315-788-1729
```

```
Alignment Scores:
Pred. No.: 0 Length: 4664
Score: 754.00 Matches: 948
```

Percent Similarity:	98.54%	Conservative:	0
Best Local Similarity:	98.54%	Mismatches:	1
Query Match:	79.45%	Indels:	14
DB:	20	Gaps:	0
US-09-671-687A-3 (1-949) x US-09-315-788-1729 (1-4664)			
Qy	1	MetSerSerGlyLeuThrSerGlnLysValThrSerProTyrTyrGluGluArgile	20
Db	230	ATGAGTTTCAGGCTTATGGAGCCAGAAAAGTCACTTCACCCCTACTGGAGAGCGGATT	289
Qy	21	PheTyrLeuLeuLeuGlnLysCysSerValThrAspLysGlnThrGlnLysLeuLys	40
Db	290	TTTTACTTGTCTTCAAGATGTCAGCTTACAGACAAACAAACAAAGCTTCCTTAA	349
Qy	41	ValProLysGlySerIleGlyGlnTyrIleGlnAspArgSerValGlyHisSerArgile	60
Db	350	GTACCGAAGGAGATATAGGACAGTATATTCAGATCGTTCTGTGGGCAATTCAGGATT	409
Qy	61	ProSerAlaLysGlyLysLysAsnGlnIleGlyLeuLysIleLeuGluGlnProHisala	80
Db	410	CCTTCGCAAAAGCGCAAGAAAATCAGATTGGATTGAAATTTAAATAATTCATAGCAACCTCATGCA	469
Qy	81	ValLeuPheValAspGlu---AspValValGluIleAsnGluLysPheThrGluLeuLeu	99
Db	470	GTTCTCTTTGTGTGATGAAGGATGTTGTAGAGATAAAATGAAAGTTTCACAGAGTTACTT	529
Qy	100	LeuAlaIleThrAsnCysGluGluArgPheSerLeuPheLysAsnArgAsnArgLeuSer	119
Db	530	TTGGCAATTCAATTTGAGAGAGGTTTACGCTGTTTAAATAACAGAAAACAGATAAGT	589
Qy	120	LysGlyLeuGlnIleAspValGlyCysProValLysValGlnLeuArgSerGlyGluGlu	139
Db	590	AAAGGCTCCAAATAGACGTGGCTGTCTGTGAAGTACAGCTGAGATCTGGGGAAGAA	649
Qy	140	LysPheProGlyValValArgPheArgGlyProLeuLeuAlaGluArgThrValSerGly	159
Db	650	AAATTTCTGTGAGTTGTACGCTTACAGAGGACCCCTGTTAGCAGAGAGACAGTCTCCGGA	709
Qy	160	IlePhePheGlyValGluLeuLeuGluGluGlyArgGlyGlnGlyPheThrAspGlyVal	179
Db	710	ATATTTCTTGAGTTGAAATTCGTAAGAGTCTGTGTCAAGTTTCACTGACGGGTG	769
Qy	180	TyrGlnGlyLysGlnLeuPheGlnCysAspGluAspCysGly---PheValAlaLeuAsp	198
Db	770	TACCAAGGAACACGCTTTTTCAGTGTGATGAAGTTGTGGCGTGTGTTGTCATTGGAC	829
Qy	199	LysLeuGluLeuIleGluAspAspThrAlaLeuGluSerAspTyrAlaGlyProGly	218
Db	830	AAGCTAGAACTCATAGAGATGATGACACTGCTTGGAAAGTGAATTACGCAAGTCTCTGGG	889
Qy	219	AspThrMetGlnValGluLeuProLeuGluIleAsnSerArgValSerLeuLysGly	238
Db	890	GACAAATGTCAGGTGCAATCTCTCTCTTTGAAATATAATCTCAGAGTTTCTTTGAA--GGT	948
Qy	239	-GlyGluThrIleGluSerGlyThrValIlePheCysAspValLeuProGlyLysGluSe	258
Db	949	TGGAGAAACAATAGNAATCGAACAGTTATATCTGTGATGTTTTCAGGAAAGAAAG	1008
Qy	258	rLeuGlyTyrPheValGlyValAspMetAspAsnProIleGlyAsnTyrAspGlyArgPh	278
Db	1009	CTTAGGATATTTGTTGTGTGGACATGATTAACCTTATGGCAACTGGGATGGAGATT	1068
Qy	278	eAspGlyVal---LeuCysSerPheAlaCysValGluSerThrIleLeuLeuHisIleAs	297
Db	1069	TGATGGAGTGAGCTTTGTAGTTTTTGGCTGTGTTGAAAGTACAAATTCATTGACATCAA	1128
Qy	297	nAspIleProGluSerValThrGlnGluArgProProLysLeuAlaPheMetSe	317
Db	1129	TGATATATCCAGAGAGTGACGACAGAGAGGAGGCTCCCAAACTTGCCCTTATGTC	1188
Qy	317	rArgGlyValGlyAspLysGlySerSerHisAsnLysProLysAlaThrGlySerTh	337

Db	1189	AAGAGGTGTTGGGAGCAAAAGGTTTCATCCAGTCATATAATAACAAAGGCTCAGAGTCTAC	1248
Qy	337	rSerAspProGlyAsnArg---ArgSerGluLeuPheTyrThrLeuAsnGlySerSerVa	356
Db	1249	CTCAGACCTCGGAATAGAAACAGATCTGAATTAATTTATACCTTAATGGTCTTCTGT	1308
Qy	356	lAspSerGlnProGlnSerLysSerLysAsnThrTyrIleAspGluValAlaGluAs	376
Db	1309	TGACTCAACACCAATCAAAATCAAAATACATGGTACATGTATGAAGTTGCAAGAAGA	1368
Qy	376	pProAlaLysSerLeuThrGluIleSerThrAspPheAspArgSerProProLeuGlu	396
Db	1369	CCCTGCAAAATCTCTTACAGAGATATCTACAGACTTTGACCGCTCTTCCACCACTCCA	1428
Qy	396	nProProValAsnSerLeuThrThrGluAsnArgPheHisSerLeuProPheSerIle	416
Db	1429	GCCTCTCTCTGTGAACCTCACTGACCAACGAGACAGATTCCACTCTTTACCATTCAGTCT	1488
Qy	416	uThrLysMetProAsnThrAsnGlySerIleGlyHisSerProLeuSerLeuSerAlaGlu	436
Db	1489	CACCAAGATGCCCAATACCAATGGAAGTATTGGCCACACAGTCCACTTTCTCTGTCAAGCCA	1548
Qy	436	nSerValMetGluGluLeuAsnThrAlaProValGlnGluSerProProLeuAlaMetPr	456
Db	1549	GTCTGTAATGGAAGAGCTAAACACATGCACCCGTCACAGAGAGTCCACCTTTGGCCATGCC	1608
Qy	456	oProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLysGluAsnProPr	476
Db	1609	TCCTGGGAACCTCACATGGTCTAGAAAGTGGGCTCATTTGGCTGAGTTAAGGAGAACCTTCC	1668
Qy	476	oPheTyrGlyValIleArgTyrIleGlyGlnProProGlyLeuAsnGluValLeuAlaGlu	496
Db	1669	TTTCTATGGGTTAATCCGTTGGATCGGTTCAGCCACCCAGGACTGGAATGAAGTCTCGCTGG	1728
Qy	496	yLeuGluLeuGluAspGluCysAlaGlyCysThrAspGlyThrPheArgGlyThrArgTy	516
Db	1729	ACTGGAACCTGGAAGATGAGTGTGCAGGCTGTACGGATGGAACCTTTACAGAGGACCTCGSTA	1788
Qy	516	rPheThrCysAlaLeuLysLysAlaLeuPheValLysLeuLysSerCysArgProAspSe	536
Db	1789	TTTCACTTGTGCCCTGAAAGAGGCGCTGTTTGTGAACTGGAAGAGCTGCAGGCTGACTTC	1848
Qy	536	rArgPheAlaSerLeuGlnProValSerAsnGlnIleGluArgCysAsnSerLeuAlaPh	556
Db	1849	TAGTGTTCATCATTTGACGCGGTTTCCATCAGATTGAGCGCTGTAACCTCTTTAGCATTT	1908
Qy	556	eGlyGlyTyrLeuSerGluValValGluGluAsnThrProProLysMetGluLysGluGlu	576
Db	1909	TGGAGGCTACTTAAGTGAAGTAGTAGAAGAAATACTCCACCAAAATGGAAGAAAGG	1968
Qy	576	yLeuGluIleMetIleGlyLysLysLysGlyIleGlnGlyHisTyrAsnSerCysTyrIle	596
Db	1969	CTTGAGATAATGATTGGGAAGAAAGGAGGATCCAGGCTCATTAACAATTTCTTTTACTT	2028
Qy	596	uAspSerThrLeuPheCysLeuPheAlaPheSerSerValLeuAspThrValLeuLeuAr	616
Db	2029	AGACTCAACCTTATTTCTGCTTATTTGCTTTAGTTCTGTTCTGGACACTGTGTTACTTAG	2088
Qy	616	gProLysGluLysAsnAspValGluTyrTyrSerGluThrGlnGluLeuLeuArgThrGlu	636
Db	2089	ACCCAAAGAAAGAACGATGTAGATAATATAGTGAACCCCAAGAGCTACTGAGGACAGA	2148
Qy	636	uIleValAsnProLeuArgIleTyrGlyTyrValCysAlaThrLysIleMetLysLeuAr	656
Db	2149	AATTTGTTAATCCTCTGAGATATATGATGTGTGTGTCACCAAAATATATGAAACTGAG	2208
Qy	656	gLysIleLeuGluLysValGluAlaAlaSerGlyPheThrSerGluGluLysAspProGlu	676
Db	2209	GAAATATCTTGAAGAGTGGAGGCTGCATCAGGATTTACCTCTGAAGAAAGATCCTGA	2268
Qy	676	uGluPheLeuAsnIleLeuPheHisIleLeuArgValGluProLeuLeuLysIleAr	696
Db	2269	GGAATTTCTTGAATATTTCTGTTTCATCATATTTTAAGGGGTAGAACCTTTGCTTAAATAAG	2328

```
Qy 696 gSerAlaGlyGlnLysValGlnAspCysTyrPheTyrGlnIlePheMetGluLysAsnG 716
Db 2329 ATCAGCGTCAAAAGGTACAGATTGTACTTCTATCAAAATTTTATCGAAAAATCA 2388
Qy 716 uLysValGlyValProThrIleGlnGlnLeuLeuGluTrpSerPheIleAsnSerAsnLe 736
Db 2389 GAAAGTTGGCGTTCCCAAAATTCAGAGTTGTAGAAATGGTCTTTTATCAACAGTAACCT 2448
Qy 736 uLysPheAlaGluAlaProSerCysLeuIleIleGlnMetProArgPheGlyLysAspPh 756
Db 2449 GAAATTTCCAGAGGCACCATCATGCTGATATTACAGATGCCTCGATTTGGAAAAAGACTT 2508
Qy 756 eLysLeuPheLysLysIlePheProSerLeuGluLeuAsnIleThrAspLeuLysGluAs 776
Db 2509 TAACTATTTAAAAAATTTTCTCTCTGAAATTAATATAACAGATTACTTACTTGAAGA 2568
Qy 776 pThrProArgGlnCysArgIleCysGlyGlyLeuAlaMetTyrGluCysArgGluCysTy 796
Db 2569 CACTCCACAGACAGTCCGCGATATGTGGAGGCTTGCAATGTATGAGTGTAGAGATGCTA 2628
Qy 796 rAspAspProAspIleSerAlaGlyLysIleLysGlnPheCysLysLysThrCysAsnThrG 816
Db 2629 CGACGATCCGACATCTCAGCTGGAAAAATCAAGCAGTTTGTAAAAACCTGCAACACTCA 2688
Qy 816 nValHisLeuHisProLysArgLeuAsnHisLysTyrAsnProValSerLeuProLysAs 836
Db 2689 AGTCACCTTCATCCGAGAGGCTGAATCATATAATATACCCAGTGTCTATCCCAAGA 2748
Qy 836 pLeuProAspTrpAspTrpArgHisGlyCysIleProCysGlnAsnMetGluLeuPheAl 856
Db 2749 CTTACCCGACTGGGACTGGAGACACGGCTGCATCCCTTGGCAGATATGAGTATTTCG 2808
Qy 856 aValLeuCysIleGluThrSerHisTyrValAlaPheValLysTyrGlyLysAspAspSe 876
Db 2809 TGTTCCTGTCATAGAAACAAGCCACTATGTTCTTGTGAAGTATGGAGAGGACGATTC 2868
Qy 876 rAlaTrpLeuPhePheAspSerMetAlaAspArgAspGlyGlnAsnGlyPheAsnIle 896
Db 2869 TGCTCGGCTCTCTTTTGACAGATCGCCGATCGGATGGTGTGTCAGAAATGGCTTCAACAT 2928
Qy 896 eProGlnValThrProCysProGluValGlyGluTyrLeuLysMetSerLeuGluAspLe 916
Db 2929 TCCTCAAGTCAACCCATGCCAGAGTAGGAGAGTACTTTGAAGATGCTCTCGAGAGACCT 2988
Qy 916 uHisSerLeuAspSerArgArgIleGlnGlyCysAlaArgArgLeuLeuCysAspAlaTy 936
Db 2989 GCATTCCTTGACTCCAGAGAAATCCAGGCTGTGCAGCAGACTGCTTTGTGTGCATA 3048
Qy 936 rMetCysMetTyrGlnSerProThrMetSerLeuTyrLys 949
Db 3049 TATGTGATGTACCAAGAGTCCCAACAAATGAGTTTGTACAAA 3088
```

## RESULT 4

```
US-09-315-788A-1729
; Sequence 1729, Application US/09315788A
; GENERAL INFORMATION:
; APPLICANT: Geating, David P.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES DERIVED FROM A
; TITLE OF INVENTION: HUMAN ASTROCYTE LIBRARY
; FILE REFERENCE: 1600.1019-002
; CURRENT APPLICATION NUMBER: US/09/315,788A
; CURRENT FILING DATE: 1999-05-21
; PRIOR FILING DATE: 09/315,788
; PRIOR FILING DATE: 1999-05-21
; PRIOR APPLICATION NUMBER: 60/086,455
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/132,067
; PRIOR FILING DATE: 1999-04-30
; NUMBER OF SEQ ID NOS: 2346
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1729
```

```
; LENGTH: 4664
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(4664)
; OTHER INFORMATION: n = A,T,C or G
US-09-315-788A-1729

Alignment Scores:
Pred. No.: 0 Length: 4664
Score: 754.00 Matches: 948
Percent Similarity: 98.54% Conservative: 0
Best Local Similarity: 98.54% Mismatches: 1
Query Match: 79.45% Indels: 14
DB: 20 Gaps: 0
```

US-09-671-687A-3 (1-949) x US-09-315-788A-1729 (1-4664)

```
Qy 1 MetSerSerGlyLeuTrpSerGlnGluLysValThrSerProTyrTrpGluGluArgIle 20
Db 230 ATGAGTTTCAAGCTTATGGAGCCAAAGAAAAGTCACCTTACCCTACTTGGGAAGAGCGATT 289
Qy 21 PheTyrLeuLeuGlnGluCysSerValThrAspLysGlnThrGlnLysLeuLys 40
Db 290 TTTTACTTCTCTTCAAGATGCGAGCTTACAGACAAACAAACACAAAAGCTCCTTAAA 349
Qy 41 ValProLysGlySerIleGlyGlnTyrIleGlnAspArgSerValGlyHisSerArgIle 60
Db 350 GTACCGAAGGAGATATAGGACAGATATATCAAGATCGTTCTGTGGGGCATTCAGGATT 409
Qy 61 ProSerAlaLysGlyLysLysAsnGlnIleGlyLeuLysIleLeuGlnProHisAla 80
Db 410 CCTTCTGCAAAAGGCAAGAAAATCAGATTGGAATTAATAATCTAGAGCAACCTCATGCA 469
Qy 81 ValLeuPheValAspGlu---AspValValGluIleAsnGlnLysPheThrGluLeuLeu 99
Db 470 GTTCTCTTTGTGATAAAGGATGTTGTAGAGATAAATGAAAAGTTTCACAGAGTTACTT 529
Qy 100 LeuAlaIleThrAsnCysGluGluArgPheSerLeuPheLysAsnArgAsnArgLeuSer 119
Db 530 TTGGCAATTAACCAATTTGTGAGGAGAGGTTTCAAGCTGTTTAAAAACAGAAACAGACAA 589
Qy 120 LysGlyLeuGlnIleAspValGlyCysProValLysValGlnLeuArgSerGlyGluGlu 139
Db 590 AAAGGCTCCAAATAGACGTGGCTGCTGTGAAAGTACAGCTGAGATCTGGGAGAA 649
Qy 140 LysPheProGlyValValArgPheArgGlyProLeuLeuAlaGluArgThrValSerGly 159
Db 650 AAATTTCTGGAGTTGTACGCTTACAGAGGACCCCTGTTAGCAGAGAGGACAGTCTCGGA 709
Qy 160 IlePhePheGlyValGluLeuLeuGluGluArgGlyGlyGlnGlyPheThrAspGlyVal 179
Db 710 ATATTTCTTTGGAGTTGAATTTGCTGGAAGAGGTCGTGGTCAAGGTTTCACTGACGGGTG 769
Qy 180 TyrGlnGlyLysGlnLeuPheGlnCysAspGluAspCysGly---PheValAlaLeuAsp 198
Db 770 TACCAAGGGAACACAGCTTTTTCAGTGTGATGAAGATTGTGGCGTGTGTTGTCATTTGGAC 829
Qy 199 LysLeuGluLeuIleGluAspAspThrAlaLeuGluSerAspTyrAlaGlyProGly 218
Db 830 AAGCTAGAACTCATGAAGATGATGACCTGCATTTGGAAGAGTATTCAGCAGGCTCTGGG 889
Qy 219 AspThrMetGlnValGluLeuProLeuGluIleAsnSerArgValSerLeuLysGly 238
Db 890 GACACATCGAGTCCGAACCTTCTCTCTTTTGGAAATAAACTCCAGAGTTTCTTTGAA-GGT 948
Qy 239 -GlyGluThrIleGluSerGlyThrValIlePheCysAspValLeuProGlyLysGluSe 258
Db 949 TGGAGAAACAATAGAAATCTGGAACAGTTATATTCTGTGATGTTTTCGAGGAAAGAAAG 1008
Qy 258 rLeuGlyTyrPheValGlyValAspMetAspAsnProIleGlyAsnTrpAspGlyArgPh 278
```

Db 1009 CTTAGGATATTTTGGTGTGGACATGGATTAACCCCTATTGGCAACTGGGATGGAAGATT 1068  
Qy eAspGlyVal---LeuCysSerPheAlaCysValGluSerThrIleLeuLeuHisIleAs 297  
Db TGATGGAGTGCAGCTTTTGTAGTTTGGCGTGTGTAAGAGTACAATTCATTGACATCAA 1128  
Qy nAspIleIleProGluSerValThrGlnGluArgProProLysLeuAlaPheMetSe 317  
Db TGATATCATCCAGAGAGTGTGACGACGAGAGAGGCGCTCCAAATTCGCTTTATGTC 1188  
Qy rArgGlyValGlyAspLysGlySerSerHisIleAsnLysProLysAlaThrGlySerTh 337  
Db AAGAGGTGTGGGACAAAGGTTCTACGTATTAATAAACCAAGGCTACAGGATCTAC 1248  
Qy rSerAspProGlyAsnArg---ArgSerGluLeuPheTyrThrLeuAsnGlySerSerVa 356  
Db CTCAGACCTGGAAATAGAACAGATCTGAATATTTATACCTTAATGGGTCTTCTGT 1308  
Qy lAspSerGlnProGlnSerLysSerLysAsnThrTrpTyrIleAspGluValAlaGluAs 376  
Db TGACTCACACCAACCAATCCAAATCAAAATAACATGATGATGAAGTTGCAGAAGA 1368  
Qy pProAlaLysSerLeuThrGluIleSerThrAspPheAspArgSerSerProLeuGlu 396  
Db CCCTGCAAAATCTCTTACAGAGATATACAGACTTTGACCGTCTCTTACACCACTCCA 1428  
Qy nProProValAsnSerLeuThrGluAsnArgPheHisSerLeuProPheSerLe 416  
Db GCCTCTCTGTGAATCTACTGACCACCGAGAACAGATTCCACTCTTTACCAATTCAGTCT 1488  
Qy uThrLysMetProAsnThrAsnGlySerIleGlyHisSerProLeuSerLeuSerAlaGl 436  
Db CACCAAGATGCCAATACCAATGAAGTATTGGCCACAGCTCCACTTTCTCTGTGAGCCA 1548  
Qy nSerValMetGluGluLeuAsnThrAlaProValGlnGluSerProProLeuAlaMetPr 456  
Db GTCTGTAAATGGAAGAGCTTAAACACATGCACCCGCTCCAAGAGAGTCCACCCCTGGCCATGCC 1608  
Qy oProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLysGluAsnProPr 476  
Db TCCTGGGAACCTACATGGTCTAGAAGTGGGCTCATTTGGCTGAAGTTAAGGAGAACCCCTCC 1668  
Qy oPheTyrGlyValIleArgTrpIleGlyGlnProProGlyLeuAsnGluValLeuAlaGl 496  
Db TTTCTATGGGTAAATCCGTGGATCGGTGAGTCCAGCCACAGACTGAATGAAGTCTCGTGG 1728  
Qy yLeuGluLeuGluAspGluCysAlaGlyCysThrAspGlyThrPheArgGlyThrArgTy 516  
Db ACTGGAACCTGGAAGATGAGTGTGCAAGGCTGTACGGATGGAACCTTCAGAGGCACTCGGTA 1788  
Qy rPheThrCysAlaLeuLysAlaLeuPheValLysLysLysSerCysArgProAspSe 536  
Db TTTACCTGTGCTCCCTGGAAGAGGCGCTGTGTGTGAACCTGAAGAGCTGCAGGCGCTGACTC 1848  
Qy rArgPheAlaSerLeuGlnProValSerAsnGlnIleGluArgCysAsnSerLeuAlaPh 556  
Db TAGGTTTGCAATATTGCAGCGGTTTCCATCAGATTGGCGCTGTAACTCTTTAGCAT 1908  
Qy eGlyGlyTyrLeuSerGluValValGluGluAsnThrProProLysMetGluLysGluGl 576  
Db TCGAGGCTACTTAAGTGAAGTAGTAGAAGAAATACTCCACCAAAATGGAAGAAAGAGG 1968  
Qy yLeuGluIleMetIleGlyLysLysGlyGlyIleGlnGlyHisTyrAsnSerCysTyrIle 596  
Db CTTGGAGATAATGATTGGGAAGAAAGAGGCAATCCAGGCTCATTAACAATTTCTTTGTTACTT 2028  
Qy uAspSerThrLeuPheCysLeuPheAlaPheSerSerValLeuAspThrValLeuLeuAr 616  
Db AGACTCAACCTTATTTCTCTTATTTGCTTTTAGTTCTGTCTTGGACACTGTGTACTTAG 2088  
Qy gProLysGluLysAsnAspValGluTyrTyrSerGluThrGlnGluLeuLeuArgThrGl 636  
Db ACCCAAGAAAGAACGATGTAGATATATATAGTGAACCCCAAGAGCTACTGTGAGACAGA 2148

Qy uIleValAsnProLeuArgIleTyrGlyTyrValCysAlaThrLysIleMetLysLeuAr 656  
Db AATTGTTAATCTCTGAGAAATATATGGATATGTGTGTGCCACAAATAATTTATGAACCTGAG 2208  
Qy gLysIleLeuGluLysValGluAlaSerGlyPheThrSerGluGluLysAspProGl 676  
Db GAAATACTCTGAAAAGGTGGAGCTGCATCAGGATTTACCTCTGAAGAAAAAGATCCTGA 2268  
Qy uGluPheLeuAsnIleLeuPheHisIleLeuArgValGluProLeuLeuLysIleAr 696  
Db GGAATCTTGAATATTTCTTTTCATCATATTTTAAAGGGTAGAACCTTTGCTTAAAAATAAG 2328  
Qy gSerAlaGlyGlnLysValGlnAspCysTyrPheTyrGlnIlePheMetGluLysAsnGl 716  
Db ATCAGCAGGTCAAAGGTACAAGATTGTACTTCTATCAATTTTATGGAAAAAATGA 2388  
Qy uLysValGlyValProThrIleGlnGlnLeuLeuGluTrpSerPheIleAsnSerAsnLe 736  
Db GAAAGTTGGCGTTCCCAATTCAGCAGTTGTAGATGGTCTTTTATCAACAGTAACCT 2448  
Qy uLysPheAlaGluAlaProSerCysLeuIleIleGlnMetProArgPheGlyLysAspPh 756  
Db GAAATTTGCAGAGGACCACATCATGCTGATTATTAGATGCCTCGATTTGGAAAAAGACTT 2508  
Qy eLysLeuPheLysLysIlePheProSerLeuGluLeuAsnIleThrAspLeuLeuGluAs 776  
Db TAAACTATTTAAAAAATTTTTTCTTCTCTGGAATTAATATACAGATTACTTCTGGAAGA 2568  
Qy pThrProArgGlnCysArgIleCysGlyGlyLeuAlaMetTyrGluCysArgGlyCysTy 796  
Db CACTCCCAGACAGTCCGATATGTGGAGGCTTCCAATGTATGATGATGATGATGATGATGAT 2628  
Qy rAspAspProAspIleSerAlaGlyLysIleLysGlnPheCysLysThrCysAsnThrGl 816  
Db GCACATCCGACATCTCAGCTGGAAAAATCAAGCAGTTTGTAAAAACCTTGCACACTCA 2688  
Qy nValHisLeuHisProLysArgLeuAsnHisLysTyrAsnProValSerLeuProLysAs 836  
Db AGTCACCTTTCATCCGAAGAGCTGAATCATATAATATTAACCCAGTGTCACTTCCCAAGA 2748  
Qy pLeuProAspTrpAspTrpArgHisGlyCysIleProCysGlnAsnMetGluLeuPheAl 856  
Db CTTACCCGACTGGGACTGGAGACACGGCTGCATCCTTGGCAGATATGAGGATTTATTC 2808  
Qy aValLeuCysIleGluThrSerHisTyrValAlaPheValLysTyrGlyLysAspAspSe 876  
Db TGTTCTCTGCATAGAAAACAAGCCACTATGTTGCTTTTGTGAAGTATGGAAGGACGATTC 2868  
Qy rAlaTrpLeuPhePheAspSerMetAlaAspArgAspGlyGlnAsnGlyPheAsnIle 896  
Db TGCTGGCTCTTCTTTTGACAGATGGCGATGGGTGTCAGAAATGGGCTTCAACAT 2928  
Qy eProGlnValThrProCysProGluValGlyGluTyrLeuLysMetSerLeuGluAspLe 916  
Db TCCTCAAGTCAACCCATGCCAGAGTAGGAGAGTACTTGAAGATGTCTCTGGAAGACCT 2988  
Qy uHisSerLeuAspSerArgArgIleGlnGlyCysAlaArgArgLeuLeuCysAspAlaTy 936  
Db GCATCTCTTGGACTCCAGAGAATCCAGGCTGTCCAGGAGCTGCTTTGTGATGCATA 3048  
Qy rMetCysMetTyrGlnSerProThrMetSerLeuTyrLys 949  
Db TATGTGCATGTACCAAGAGTCCAAATGAGTTTGTACAAA 3088

## RESULT 5

US-09-783-514-1729  
; Sequence 1729, Application US/09783514  
; GENERAL INFORMATION:  
; APPLICANT: Gearing, David P.  
; APPLICANT: Holtzman, Douglas A.  
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES DERIVED FROM A  
; TITLE OF INVENTION: HUMAN ASTROCYTE LIBRARY

FILE REFERENCE: 1600.1019-002  
 CURRENT APPLICATION NUMBER: US/09/783,514  
 CURRENT FILING DATE: 2001-02-14  
 PRIOR APPLICATION NUMBER: 09/315,788  
 PRIOR FILING DATE: 1999-05-21  
 PRIOR APPLICATION NUMBER: 60/086,455  
 PRIOR FILING DATE: 1998-05-22  
 PRIOR APPLICATION NUMBER: 60/132,067  
 PRIOR FILING DATE: 1999-04-30  
 NUMBER OF SEQ ID NOS: 2346  
 SOFTWARE: FastSeq for Windows Version 4.0  
 SEQ ID NO 1729  
 LENGTH: 4664  
 TYPE: DNA  
 ORGANISM: Homo sapien  
 FEATURE:  
 NAME/KEY: misc\_feature  
 LOCATION: (1)...(4664)  
 OTHER INFORMATION: n = A,T,C or G  
 US-09-783-514-1729

Alignment Scores:  
 Pred. No.: 0 Length: 4664  
 Score: 754.00 Matches: 948  
 Percent Similarity: 98.54% Conservative: 0  
 Best Local Similarity: 98.54% Mismatches: 1  
 Query Match: 79.45% Indels: 14  
 DB: 34 Gaps: 0

US-09-671-687A-3 (1-949) x US-09-783-514-1729 (1-4664)

QY	1	MetSerGlyLeuTyrSerGlnGluLysValThrSerProTyrTyrTyrGluGluArgile	20
DB	230	ATGAGTTTCAGCTTATGAGGCCAAGAAAAGTCACTTCCCTACTTGGGAAGCGGATT	289
QY	21	PheTyrLeuLeuLeuGlnGluCysSerValThrAspLysGlnThrGlnLysLeuLeuLys	40
DB	290	TTTTTACTTGCTTCTTCAAGATGTCAGCGTTACAGACAAACAAACAAAGCTCTTTAA	349
QY	41	ValProLysGlySerIleGlyGlnTyrIleGlnAspArgSerValGlyHisSerArgile	60
DB	350	GTACCGAAGGGAAGTATAGGACAGTATATTCAAGATCGTTCTGTGGGCAATTCAGGATT	409
QY	61	ProSerAlaLysGlyLysLysAsnGlnIleGlyLeuLysIleLeuGluGlnProHisAla	80
DB	410	CCCTTCTGCAAAAGGCAAGAAAATCAGATTGGATTAAAAATTTCTAGAGCAACTCATGCA	469
QY	81	ValLeuPheValAspGlu---AspValValGluIleAsnGluLysPheThrGluLeuLeu	99
DB	470	GTTCCTCTTTGATGAAAAGGATGTTGTAGAGATATATGAAAAGTTACAGAGTTACTT	529
QY	100	LeuAlaIleThrAsnCysGluGluArgPheSerLeuPheLysAsnArgLeuSer	119
DB	530	TTGGCAATTACCAATTGTGAGGAGAGGTTTCAGCTGTTTAAAAAACAGAAACAGACTAAGT	589
QY	120	LysGlyLeuGlnIleAspValGlyCysProValLysValGlnLeuArgSerGlyGluGlu	139
DB	590	AAAGGCCCTCCAAATAGACGTGGGCTGTCCTGTGAAAAGTACAGCTGAGATCTGGGGAAGA	649
QY	140	LysPheProGlyValValArgPheArgGlyProLeuLeuAlaGluArgThrValSerGly	159
DB	650	AAATTTCTTGAGTTGTACGTTACAGAGACCCCTTTTAGCAGAGAGGACAGTCTCCGGA	709
QY	160	IlePhePheGlyValGluLeuLeuGluGluArgGlyGlnGlyPheThrAspGlyVal	179
DB	710	ATATTTCTTGAGTTGAAATGCTGGAAGAAGTCTGTGTGTCAGGTTTCAAGTTTCACTGACGGG	769
QY	180	TyrGlnGlyLysGlnLeuPheGlnCysAspGluAspCysGly---PheValAlaLeuAsp	198
DB	770	TACCAAGGGAACACGCTTTTTCAGTGTGATGAAGATTGTGGCGTGTGTTGTCATTGGAC	829
QY	199	LysLeuGluLeuIleGluAspAspThrAlaLeuGluSerAspTyrAlaGlyProGly	218

DB	830	AAGTAGAACTCATAGAAGATGATGACACTGCTGAAAGTGAATTACGACGCTCTGGG	889
QY	219	AspThrMetGlnValGluLeuProLeuGluIleAsnSerArgValSerLeuLysGly	238
DB	890	GACACAATGACAGTCAAACTTCTCTTTTGGAAATAAACTCCAGAGCTTCTTTGAA--GGT	948
QY	239	-GlyGluThrIleGluSerGlyThrValIlePheCysAspValLeuProGlyLysGluSe	258
DB	949	TGGAGAAAACAATAGAAATCTGGAAACAGTTATATCTGTGATGTTTTCGCCAGGAAAAG	1008
QY	258	rLeuGlyTyrPheValGlyValAspMetAspAsnProIleGlyAsnTyrAspGlyArgph	278
DB	1009	CTTAGGATATTTTGTGTGTGGACATGATTAACCTATTGGCACTGGATGGAAGATT	1068
QY	278	eAspGlyVal---LeuCysSerPheAlaCysValGluSerThrIleLeuLeuHisIleAs	297
DB	1069	TGATGGAGTGCAGCTTTGTAGTTTTCGGTGTGTGTGAAGTACAAATCTATTGCACATCA	1128
QY	297	nAspIleIleProGluSerValThrGlnGluArgProProLysLeuAlaPheMetSe	317
DB	1129	TGATATCATCCAGAGAGTGTGACGCAAGAAAGAGCGCTCCCAAACTTGCTTTATGTC	1188
QY	317	rArgGlyValGlyAspLysGlySerSerSerHisAsnLysProLysAlaThrGlySerTh	337
DB	1189	AAAGAGTGTGGGGAACAAAGGTTTCATCCAGTCAATAATAACCAAGGCTACAGGATCTAC	1248
QY	337	rSerAspProGlyAsnArg---ArgSerGluLeuPheTyrThrLeuAsnGlySerSerVa	356
DB	1249	CTCAGACCTCGGAATAGAAACAGATCTGAATTTATACCTTAATTTGGGCTTCTGT	1308
QY	356	lAspSerGlnProGlnSerLysSerLysAsnThrTyrTyrIleAspGluValAlaGluAs	376
DB	1309	TGACTCACACCAACCAATCCAATCAAAAATAATCATGGTACATTTGATGAAGTTGCAGAAGA	1368
QY	376	pProAlaLysSerLeuThrGluIleSerThrAspPheAspArgSerSerProLeuGlu	396
DB	1369	CCCTGCAAAATCTTTACAGAGATATCTACAGACTTTGACCGTCTTTCACACCACTCCA	1428
QY	396	nProProValAsnSerLeuThrThrGluAsnArgPheHisSerLeuProPheSerLe	416
DB	1429	GGCTCTCTCTGTGAATCACTGACCACCGAGAACAGATTCCACTCTTTACCATTCAGTCT	1488
QY	416	uThrLysMetProAsnThrAsnGlySerIleGlyHisSerProLeuSerLeuSerAlaGl	436
DB	1489	GACCAAGATGCCCAATACCAATGGAAGTATTGGCCACAGCTCCACTTTCTCTGTGACCCCA	1548
QY	436	nSerValMetGluGluLeuAsnThrAlaProValGlnGluSerProLeuAlaMetPr	456
DB	1549	GTCTGTAATGGAAGAGCTAAACACTGCACCCGTCACAGAGAGTCCACCTTGGCCCATGCC	1608
QY	456	oProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLysGluAsnProPr	476
DB	1609	TCCTGGGAATCACAATGGTCTAGAAGTGGGCTCATTTGGCTGAAGTTAAGAGAACCTCC	1668
QY	476	oPheTyrGlyValIleArgTyrIleGlyGlnProProGlyLeuAsnGluValLeuAlaGl	496
DB	1669	TTTCTATGGGGTAATCCGTTGGATCGGTCAACCACAGAGTCAAGTGAAGTGTCTCGCTGG	1728
QY	496	yLeuGluLeuGluAspGluCysAlaGlyCysThrAspGlyThrPheArgGlyThrArgTy	516
DB	1729	ACTGGAACCTGGAAGATGAGTGTGACGGCTGACGGATGGAACCTTCAGAGGCACTCGGTA	1788
QY	516	rPheThrCysAlaLeuLysLysAlaLeuPheValLysLeuLysSerCysArgProAspSe	536
DB	1789	TTTACCTGTGCGCTGAAGAAGCGCTGTTGTGTGAATACTGAAGAGCTGCAGGCGCTGACTC	1848
QY	536	rArgPheAlaSerLeuGlnProValSerAsnGlnIleGluArgCysAsnSerLeuAlaPh	556
DB	1849	TAGTTTTCATCATTCGACCGCGTTTCCAATCAGATTGACGCTGTAACTCTTTTAGCATT	1908
QY	556	eGlyGlyTyrLeuSerGluValValGluGluAsnThrProProLysMetGluLysGluGl	576
DB	1909	TGGAGGCTACTTAAAGTAGTAGAAGAAAATACTCCACCAAAAATGGAAGAAAGG	1968

```
Qy 576 YLeuGluileMetIleGlyLysLysGlyGlyIleGlnGlyHisTyrAsnSerCysTyrzle 596
Db 1969 CTTGAGATAATGATTGGGAAGAAAGGCAATCCAGGCTCATTAACAATTCCTGTACTT 2028

Qy 596 uAspSerThrLeuPheCysLeuPheAlaPheSerSerValLeuAspThrValLeuLeuAr 616
Db 2029 AGACTCAACCTTATTCTCTCTATTTCCTTTTAGTTCTGTCTCGGACACTGTGTACTTAG 2088

Qy 616 qProLysGluLysAsnAspValGluTyrSerGluThrGlnGluLeuLeuArThrGln 636
Db 2089 ACCCAAGAAAGAACGATGTAGATATATATAGTGAACCCCAAGAGCTACTGAGGACAGA 2148

Qy 636 uileValAsnProLeuArGileTyrGlyTyrValCysAlaThrLysIleMetLysLeuAr 656
Db 2149 AATTGTTAATCTCTGAGAAATATATGATATGTTGTGTCACCAAAATATATCAACTGAG 2208

Qy 656 gLysIleLeuGluLysValGluAlaSerGlyPheThrSerGluGluLysAspProGln 676
Db 2209 GAAATATCTTGAAAAGGTGGAGGCTGCATCAGGATTTACCTCTGAAGAAAAGATCCTGA 2268

Qy 676 uGluPheLeuAsnIleLeuPheHisTyrLeuArgValGluProLeuLeuLysIleAr 696
Db 2269 GGAATCTTGTAATATCTCTTTCATCATATTTTAAGGGTAGAACCTTTGCTAAAAATAAG 2328

Qy 696 gSerAlaGlyGlnLysValGlnAspCysTyrPheTyrGlnIlePheMetGluLysAsnGln 716
Db 2329 ATCAGCAGCTCAAAAGGTACAAGATTGTTCTCTATCAATTTTATGGAATAAATGA 2388

Qy 716 uLysValGlyValProThrIleGlnLeuLeuGluTrpSerPheIleAsnSerAsnLe 736
Db 2389 GAAAGTTGGCGTTCACCAATTCAGCAGTGTGTAGATGGTCTTTTATCAACAGTAACCT 2448

Qy 736 uLysPheAlaGluAlaProSerCysLeuIleIleGlnMetProArgPheGlyLysAspPh 756
Db 2449 GAAATTTTCAGAGGACCATCATCATGCTGATTAATCAGATGCTCGATTGGAAAAGACTT 2508

Qy 756 eLysLeuPheLysIlePheProSerLeuGluLeuAsnIleThrAspLeuLeuGluAs 776
Db 2509 TAACTATTAAAAAATTTTCTCTCTGGAATTAATATAACAGATTACTTACTGAGA 2568

Qy 776 pThrProArgGlnCysArGileCysGlyGlyLeuAlaMetTyrGluCysArgGluCysGly 796
Db 2569 CACTCCCAAGACAGTCCCGATATGTGGAGGCTTCAATGTATGATGTTAGAGATGCTA 2628

Qy 796 rAspAspProAspIleSerAlaGlyLysIleLysGlnPheCysLysThrCysAsnThrGln 816
Db 2629 CGAGCATCCGACATCTCAGCTGGAAAAATCAAGCAGTGTGTGTAATAACCTGCAACTCA 2688

Qy 816 nValHisLeuHisProLysArgLeuAsnHisLysTyrAsnProValSerLeuProLysAs 836
Db 2689 AGTCACCTTCATCCGAAGAGCTGAATCATAAATATAACCCAGTGTCACTTCCCAAGA 2748

Qy 836 pLeuProAspTrpAspTrpArgHisGlyCysIleProCysGlnAsnMetGluLeuPheAl 856
Db 2749 CTTACCCGACTGGGACTGGAGACACGGCTGCATCCCTGGCCAGATATGAGATTATTGC 2808

Qy 856 aValLeuCysIleGluThrSerHisTyrValAlaPheValLysTyrGlyLysAspAspSe 876
Db 2809 TGTCTCTGTCATAGAAAACAAGCCACTATGTGTCTTTTGTGAAGTATGGAAGGACGATTC 2868

Qy 876 rAlaTrpLeuPhePheAspSerMetAlaAspArgAspGlyGlnAsnGlyPheAsnIle 896
Db 2869 TGCCTGGCTCTCTTTTGACAGATGGCCGATCGGATGGTGTGTCAGATGGCTTCAACAT 2928

Qy 896 eProGlnValThrProCysProGluValGlyGluTyrLeuLysMetSerLeuGluAspLe 916
Db 2929 TCCTCACTCACCCATGCCAGAGTAGGAGAGTACTTGAAGATGTCTCTGGAAGACCT 2988

Qy 916 uHisSerLeuAspSerArgArgIleGlnGlyCysAlaArgArgLeuLeuCysAspAlaArg 936
Db 2989 GCATTCCTTGAGTCTCAGAGAAATCCAGGCTGTGCAAGAGCTGCTTTGTGTGTCATA 3048
```

```
Qy 936 rMetCysMetTyrGlnSerProThrMetSerLeuTyrLys 949
Db 3049 TATGTGCATGTACCAAGAGTCCAAACAATGAGTTTGTACAAA 3088
```

## RESULT 6

```
US-09-783-514A-1729
; Sequence 1729, Application US/09783514A
; GENERAL INFORMATION:
; APPLICANT: Gearing, David P.
; APPLICANT: Holtzman, Douglas A.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES DERIVED FROM A
; TITLE OF INVENTION: HUMAN ASTROCYTE LIBRARY
; FILE REFERENCE: 1600.1019-002
; CURRENT APPLICATION NUMBER: US/09/783,514A
; PRIOR FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: 09/315,788
; PRIOR FILING DATE: 1999-05-21
; PRIOR APPLICATION NUMBER: 60/086,455
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/132,067
; PRIOR FILING DATE: 1999-04-30
; NUMBER OF SEQ ID NOS: 2346
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1729
; LENGTH: 4664
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURES:
; NAME/KEY: misc_feature
; LOCATION: (1)...(4664)
; OTHER INFORMATION: n = A,T,C or G
US-09-783-514A-1729
```

```
Alignment Scores:
Pred. No.: 0 Length: 4664
Score: 754.00 Matches: 948
Percent Similarity: 98.54% Conservative: 0
Best Local Similarity: 98.54% Mismatches: 1
Query Match: 79.45% Indels: 14
DB: 34 Gaps: 0
```

US-09-671-687A-3 (1-949) x US-09-783-514A-1729 (1-4664)

```
Qy 1 MetSerSerGlyLeuTrpSerGlnGluLysValThrSerProTyrTrpGluArgIle 20
Db 230 ATGAGTTTCAGGCTTATGAGGCCAAGAAAAGTCACTTCACCTACTCTGGGAGCGGATT 289

Qy 21 PheTyrLeuLeuLeuGlnGluCysSerValThrAspLysGlnThrGlnLysLeuLys 40
Db 290 TTTTACTTGCTTCTTCAAGATGCGCTTACAGCAACAACAACAACAAGCTCTTAAA 349

Qy 41 ValProLysGlySerIleGlyGlnTyrIleGlnAspArgSerValGlyHisSerArgIle 60
Db 350 GTACCGAAGGGAAGTATAGGACAGTATATTCAAGATCGTTCTGTGGGCATTCAAGATT 409

Qy 61 ProSerAlaLysGlyLysLysAsnGlnIleGlyLeuLysIleLeuGluGlnProHisAla 80
Db 410 CCTTCTGCAAGGAAGCAAGAAAATCAGATTGGATTAAAAAATTTCTAGAGCAACCTCATGCA 469

Qy 81 ValLeuPheValAspGluLysValValValGluIleAsnGluLysPheThrGluLeuLeu 99
Db 470 GTTCTCTTTTGTATGAAAAGAGTGTGTAGAGATAAATGAAAAGTTTACAGAGTTACTTT 529

Qy 100 LeuAlaIleThrAsnCysGluGluArgPheSerLeuPheLysAsnArgAsnArgLeuSer 119
Db 530 TTGGCAATTACCAATTGTGAGGAGAGGTTTCAGCTGTTTAAAAACAGAAACAGACTAAGT 589

Qy 120 LysGlyLeuGlnIleAspValGlyCysProValLysValGlnLeuArgSerGlyGluGlu 139
Db 590 AAAGCCCTCCAAATAGACGTGGCTGTCTCTGTGAAAGTACAGCTCAGATCTGGGGAAGAA 649

Qy 140 LysPheProGlyValValArgPheArgGlyProLeuLeuAlaGluArgThrValSerGly 159
```



Db 650 AAAATTCCTGGAGTTGTACGCTTCAGAGGACCCCTCTTAGCAGAGAGGACAGCTCCGGA 709  
Qy 160 IlePhePheGlyValGluLeuLeuGluGlyArgGlyGlnGlyPheThrAspGlyVal 179  
Db 710 ATATTCCTTGGAGTTGAATGCTGGAAGAAGTCGTGGTCAAGGTTTCACAGCGGGTG 769  
Qy 180 TyrGlnGlyLysGlnLeuPheGlnCysAspGluAspCysGly---PheValAlaLeuAsp 198  
Db 770 TACCAGGGAACACAGCTTTTCAGTGTGATGAAGATTGTGGCGTGTGTTGTCATTGGAC 829  
Qy 199 LysLeuGluLeuIleGluAspAspThrAlaLeuGluSerAspThrAlaGlyProGly 218  
Db 830 AAGCTAGAACTCATAGAAGATGATGACACTGCATTGGAAAGTCAATTACGAGGCTCTGG 889  
Qy 219 AspThrMetGlnValGluLeuProProLeuGluIleAsnSerArgValSerLeuLysGly 238  
Db 890 GACCAATGCGAGTGGAACTTCCTCTCTTGGAAATAAACTCCAGAGTTCTTTGAA-GGT 948  
Qy 239 -GlyGluThrIleGluSerGlyThrValIlePheCysAspValLeuProGlyLysGluSe 258  
Db 949 TGGAGAACATAGAACTGGAAACAGTTATATCTGTGATGTTTCCAGGAAAGAAAG 1008  
Qy 258 rLeuGlyTyrPheValGlyValAspMetAspAsnProIleGlyAsnTrpAspGlyArgPh 278  
Db 1009 CTTAGGATATTTTGTGGTGGACATGCATAACTTATGGCAACTGGGATGGAAGATT 1068  
Qy 278 eAspGlyVal---LeuCysSerPheAlaCysValGluSerThrIleLeuLeuHisIleAs 297  
Db 1069 TGATGGAGTCGAGCTTTGTAGTTTTCGCGTGTGTGAAGTACAATTCATTCACATCAA 1128  
Qy 297 nAspIleIleProGluSerValThrGlnGluArgArgProProLysLeuAlaPheMetSe 317  
Db 1129 TGATATATCCAGAGAGTGTGACGAGGAAGAGGCGCTCCCAAACTTGCCCTTTATGTC 1188  
Qy 317 rArgGlyValGlyAspLysGlySerSerHisAsnLysProLysAlaThrGlySerTh 337  
Db 1189 AAGAGTGTGGGGCAAAAGTTTCATCAGTCATATAAACAACCAAGGCTACAGGACTAC 1248  
Qy 337 rSerAspProGlyAsnArg---ArgSerGluLeuPheThrLeuAsnGlySerSerVa 356  
Db 1249 CTCAGACCTCGAAATAGAAACAGATCTCAATATTTTATACCTTAAATGGGTCTTCTGT 1308  
Qy 356 lAspSerGlnProGlnSerLysSerLysAsnThrTrpTyrIleAspGluValAlaGluAs 376  
Db 1309 TGACTCAACACCAATCAAAATCAAAATAATCATGGTACATTTGATGAAGTTGCAGAGA 1368  
Qy 376 pProAlaLysSerLeuThrGluIleSerThrAspPheAspArgSerSerProLeuGlu 396  
Db 1369 CCCTGCAAAATCTTTACAGAGATATCTACAGACTTTGACCGTCTTTCACCACTCCA 1428  
Qy 396 nProProValAsnSerLeuThrThrGluAsnArgPheHisSerLeuProPheSerLe 416  
Db 1429 GCCTCCTCTGTGAACCTACTGACCAACCCAGAACAGATCCACTCTTTACCAATTCAGTCT 1488  
Qy 416 uThrLysMetProAsnThrAsnGlySerIleGlyHisSerProLeuSerLeuSerIle 436  
Db 1489 CACCAGAGTGCCCAATACAAATGGAAGTATGGCCACAGTCACACTTCTCTGTGAGCCCA 1548  
Qy 436 nSerValMetGluGluLeuAsnThrAlaProValGlnGluSerProProLeuAlaMetPr 456  
Db 1549 GTCTGTATGGAAGAGCTAAACATGCAACCGTCCAAAGAGATCCACCTTGGCCATGCC 1608  
Qy 456 oProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLysGluAsnProPr 476  
Db 1609 TCCTGGGAACCTCACATGGTCTAGAGTGGGCTCATTTGGCTGAAGTTAAGGAGAACCTCC 1668  
Qy 476 oPheTyrGlyValIleArgTrpIleGlyGlnProProGlyLeuAsnGluValLeuIle 496  
Db 1669 TTTCTATGGGGTAATCCGTTGATCGGTGAGCCACAGGACTGAATGAAGTGTCTCGCTGG 1728  
Qy 496 yLeuGluLeuGluAspGluCysAlaGlyCysThrAspGlyThrPheArgGlyThrArgTy 516  
Db 1729 ACTGAACTGGAGATGAGTGTGAGGCTGTACGGATGAACTTCAGAGGCACTCGGTA 1788

Qy 516 rPheThrCysAlaLeuLysLysAlaLeuPheValLysLeuLysSerCysArgProAspSe 536  
Db 1789 TTTACCTCTGCCTGAAGAGCGCTGTTTGTGAAACTGAAGAGCTGCAGGCTGACTC 1848  
Qy 536 rArgPheAlaSerLeuGlnProValSerAsnGlnIleGluArgCysAsnSerLeuAlaPh 556  
Db 1849 TAGGTTTGCATCATTCAGCGGCTTCCAATCAGATTAGCGCTGTAACCTCTTAGCAT 1908  
Qy 556 eGlyGlyTyrLeuSerGluValValGluGluAsnThrProProLysMetGluLysGluGl 576  
Db 1909 TGGAGGCTACTTAAGTGAAGTAGTAGAAGAAATACTCCACCAAAATGGAAGAAAG 1968  
Qy 576 yLeuGluIleMetIleGlyLysLysGlyIleGlnGlyHisTyrAsnSerCysTyrLe 596  
Db 1969 CTTGGAGATATGATTGGGAAGAAAGAGCATCCAGGTCATTACAAATCTTGTACTT 2028  
Qy 596 uAspSerThrLeuPheCysLeuPheAlaPheSerSerValLeuAspThrValLeuLeuAr 616  
Db 2029 AGACTCAACTTATCTGCTTATTTGCTTTAGTTCTGTTCTGGACACTGTGTACTTAG 2088  
Qy 616 gProLysGluLysAsnAspValGluTyrTyrSerGluThrGlnGluLeuLeuArgThrGl 636  
Db 2089 ACCCAAGAAAGACAGATGATATATATAGTGAACCCAAAGAGCTACTGAGGACAGA 2148  
Qy 636 uIleValAsnProLeuArgIleTyrGlyTyrValCysAlaThrLysIleMetLysIleAr 656  
Db 2149 AATTGTTAATCTCTCAGAAATATATGATATGTTGTGCCCAAAATAATTAATGAAC 2208  
Qy 656 gLysIleLeuGluLysValGluAlaAlaSerGlyPheThrSerGluLysAspProGl 676  
Db 2209 GAAATACCTTGAAGAGTGGAGGCTGCATCAGGATTTACCTCGAAGAAAGATCTCGA 2268  
Qy 676 uGluPheLeuAsnIleLeuPheHisIleLeuArgValGluProLeuLeuLysIleAr 696  
Db 2269 GGAATCTTGAATATCTGTTTCATCATATTTTAAAGGGTAGAACCTTTGCTTAAATA 2328  
Qy 696 gSerAlaGlyGlnLysValGlnAspCysTyrPheTyrGlnIlePheMetGluLysAsnGl 716  
Db 2329 ATCAGAGGTCAAAAGGTACAGATTGTTACTTCTATCAAAATTTTATGGAAGAAATA 2388  
Qy 716 uLysValGlyValProThrIleGlnGlnLeuLeuGluTrpSerPheIleAsnSerAsnLe 736  
Db 2389 GAAAGTTGGCTTCCACCAATTCAGCAGTTGTTAGATGCTCTTTTATCAACAGTAAC 2448  
Qy 736 uLysPheAlaGluAlaProSerCysLeuIleIleGlnMetProArgPheGlyLysAspPh 756  
Db 2449 GAAATTTGCAGAGGCAACCATCATGTTCTGATTAATTCAGATGCCCTCGATTGGA 2508  
Qy 756 eLysLeuPheLysLysIlePheProSerLeuGluLeuAsnIleThrAspLeuLeuAs 776  
Db 2509 TAAACTATTTTAAAAAATTTTCTCTCTCGAAATTAATATAACAGATTACTTTGAAGA 2568  
Qy 776 pThrProArgGlnCysArgIleCysGlyLeuAlaMetTyrGluCysArgGluCysTy 796  
Db 2569 CACTCCAGACAGTGCAGATATGTGAGGCGCTTGCATGATGATGAGCTGAGAGATGCTA 2628  
Qy 796 rAspAspProAspIleSerAlaGlyLysIleLysGlnPheCysLysThrCysAsnThrGl 816  
Db 2629 CGACGATCCGGACATCTCAGCTGGAATAATCAAGCAGTTTGTAAAAACCTGCAACTCA 2688  
Qy 816 nValHisLeuHisProLysArgLeuAsnHisLysTyrAsnProValSerLeuProLysAs 836  
Db 2689 AGTCCACCTTCATCCGAAGAGGCTGAATCAATAATATAACCCAGTGCATCTCCCAAGA 2748  
Qy 836 pLeuProAspTrpAspTrpArgHisGlyCysIleProCysGlnAsnMetGluLeuPheAl 856  
Db 2749 CTTACCCGACTGGGACTGGAGACCGGCTGCATCCCTTGCAGAAATATGAGATTATTGC 2808  
Qy 856 aValLeuCysIleGluThrSerHisTyrValAlaPheValLysTyrGlyLysAspAspSe 876  
Db 2809 TGTTCCTGATAGAAACCAAGCCACTATGTTGCTTTTGTGAAGTATGGAAGGACGATTC 2868



```
Qy 876 rAlaTrpLeuPhePheAspSerMetAlaAspArgAspGlyGlyGlnAenGlyPheAsnI1 896
Db 2869 TGCCTGGCTCTTCTTGACAGCATGCCGATCGGATGGTGTGTCAGATGGCTTCAACAT 2928
Qy 896 eProGlnValThrProCysProGluValGlyGluTyrLeuLysMetSerLeuGluAspLe 916
Db 2929 TCCTCAAGTCACCCCATGCCAGAGTAGTAGGAGTACTTGAAGATGTCTCTGGAAGACCT 2988
Qy 916 uHisSerLeuAspSerArgArgGileGlnGlyCysAlaArgArgLeuLeuCysAspAlaty 936
Db 2989 GCATTCTTGACCTCAGAGATCCAGGCTGTGCAGAGACTGCTTTGTGATGCATA 3048
Qy 936 rMetCysMetTyrGlnSerProThrMetSerLeuTyrLys 949
Db 3049 TATGTGATGTACCAAGAGTCCCAACATGAGTTGTACAAA 3088

RESULT 7
US-09-396-087-4309
; Sequence 4309, Application US/09396087
; GENERAL INFORMATION:
; APPLICANT: Geating, David P.
; APPLICANT: Holtzman, Douglas A.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES DERIVED FROM A
; TITLE OF INVENTION: PULMONARY ARTERY SMOOTH MUSCLE LIBRARY
; FILE REFERENCE: MLN98-39pm
; CURRENT APPLICATION NUMBER: US/09/396,087
; CURRENT FILING DATE: 1999-09-14
; EARLIER APPLICATION NUMBER: 60/100,260
; EARLIER FILING DATE: 1998-09-14
; EARLIER APPLICATION NUMBER: 60/107,226
; EARLIER FILING DATE: 1998-11-05
; EARLIER APPLICATION NUMBER: 60/131,810
; EARLIER FILING DATE: 1999-04-30
; NUMBER OF SEQ ID NOS: 5220
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4309
; LENGTH: 4668
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(4668)
; OTHER INFORMATION: n = A,T,C or G
US-09-396-087-4309

Alignment Scores:
Pred. No.: 0 Length: 4668
Score: 754.00 Matches: 948
Percent Similarity: 98.54% Conservative: 0
Best Local Similarity: 98.54% Mismatches: 1
Query Match: 79.45% Indels: 14
DB: 21 Gaps: 0

US-09-671-687A-3 (1-949) x US-09-396-087-4309 (1-4668)

Qy 1 MetSerGlyLeuTrpSerGlnGluLysValThrSerProTyrTrpGluArgile 20
Db 230 ATGAGTTTCTTCTTCAAGATGAGCGTTACAGCAACCAACCAAGCTCTTAAA 289
Qy 21 PheTyrLeuLeuGlnGluCysSerValThrAspLysGlnThrGlnLysLeuLys 40
Db 290 TTTTACTTCTTCTTCAAGATGAGCGTTACAGCAACCAACCAAGCTCTTAAA 349
Qy 41 ValProLysGlySerileGlyGlnTyrileGlnAspArgSerValGlyHisSerArgile 60
Db 350 GTACCGAAGGAGGATATAGGACAGTATATTCAAGATGTTCTGTGGGCAATCAAGGATT 409
Qy 61 ProSerAlaLysGlyLysLysAsnGlnileGlyLeuLysLysLeuGluGlnProHisala 80
Db 410 CCTTCTGCAAAAGGCAAGAAAAATCAGATTGGATTAAAAATTTCTAGACCAACCTCATGCA 469
Qy 81 ValLeuPheValAspGlu---AspValValGluLeuLeuLysPheThrGluLeuLeu 99
```

```
Db 470 GTTCTCTTTTGTGATGAAAGGATGTTGTAGAGATAAAATGAAAGTTCACAGAGTTACTT 529
Qy 100 LeuAlaileThrAsnCysGluGluArgPheSerLeuPheLysAsnArgAsnArgLeuSer 119
Db 530 TTGGCAATTTACCAATTTGTAGAGAGGTTTACGCTGTTTAAAAACAGAAACAGACTAAGT 589
Qy 120 LysGlyLeuGlnIleAspValGlyCysProValLysValGlnLeuArgSerGlyGluGlu 139
Db 590 AAAGCCCTCCAAATAGACGTGGCTGCTCTGTGAAAGTACAGTCTGAGATCTCGGGAAGA 649
Qy 140 LysPheProGlyValValArgPheArgGlyProLeuLeuAlaGluArgThrValSerGly 159
Db 650 AAATTTCTGGAGTTGTACGCTTACAGAGACCCCTGTTAGCAGAGAGGACAGTCTCCGA 709
Qy 160 IlePhePheGlyValGluLeuLeuGluArgGlyGlnGlyGlnPheThrAspGlyVal 179
Db 710 ATATTCTTTGGAGTTGAATTTGCTGGAAGAGGTCGTGGTCAAGGTTTCTACTACGGGGT 769
Qy 180 TyrGlnGlyLysGlnLeuPheGlnCysAspGluAspCysGly---PheValAlaLeuAsp 198
Db 770 TACCAAGGGAACACAGCTTTTTCAGTGTGATGAGAGTTTGGCGTGTGTTGTCATTTGAC 829
Qy 199 LysLeuGluLeuIleGluAspAspThrAlaLeuGluSerAspTyrAlaGlyProGly 218
Db 830 AAGCTAGAACTCATAGAAGATGATGACACTGCTTGGAAAGTATTCACGAGGTCTCTGG 889
Qy 219 AspThrMetGlnValGluLeuProLeuGluIleAsnSerArgValSerLeuLysGly 238
Db 890 GACACAAATGCAAGGTCGAACTTCTCTTTGGAAATAAACTCCAGAGTTTCTTTGAA-GGT 948
Qy 239 -GlyGluThrIleGluSerGlyThrValIlePheCysAspValLeuProGlyLysGlu 258
Db 949 TGGGAAACAATAGAACTCTGGACAGTTATTTCTGTGATGTTTTCAGGAAAGAAAG 1008
Qy 258 rLeuGlyTyrPheValGlyValAspMetAspAsnProIleGlyAsnTrpAspGlyArgPh 278
Db 1009 CTTAGGATATTTTGTGTTGGTGGACATGGATAACCTATTGGCAACTGGGATGGAAGATT 1068
Qy 278 eAspGlyVal---LeuCysSerPheAlaCysValGluSerThrIleLeuLeuHisIleAs 297
Db 1069 TGATGGAGTGCAGCTTTTGTAGTTTTCGTGTTTGGAAAGTACAAATTCATTGACATCAA 1128
Qy 297 nAspIleileProGluSerValThrGlnGluArgArgProProLysLeuAlaPheMetSe 317
Db 1129 TGATATCATCCAGAGAGTGTACCCAGGAAGAGGCTCCCAACTTGGCTTTATGTATGTC 1188
Qy 317 rArgGlyValGlyAspLysGlySerSerHisAsnLysProLysAlaThrGlySerTh 337
Db 1189 AAGAGTGTGGGACAAAGTTTCATCCAGTCATAATAAACCAAGGCTACAGGATCTAC 1248
Qy 337 rSerAspProGlyAsnArg---ArgSerGluLeuPheTyrThrLeuAsnGlySerSerVa 356
Db 1249 CTCAGACCTCGAAATAGAAACAGATCTGAATTTATTTATACCTTAAATGGTCTTCTGT 1308
Qy 356 lAspSerGlnProGlnSerLysSerLysAsnThrTrpTyrileAspGluValAlaGluAs 376
Db 1309 TGACTCACACACCAATCAAAATCAAAATAACATGGTACATGATGAAGTTGCGAAGA 1368
Qy 376 pProAlaLysSerLeuThrGluIleSerThrAspPheAspArgSerSerProLeuGlu 396
Db 1369 CCTTGCAAAATCTTTACAGAGATATCTACAGACTTTCACGCTTCTTCCACCACTCCA 1428
Qy 396 nProProValAsnSerLeuThrGluAsnArgPheHisSerLeuProPheSerle 416
Db 1429 GCCTCTCTCTGTGAATCACTGACCACCGAGAACAGATTCCACTCTTTACCATTCAGTCT 1488
Qy 416 uThrLysMetProLanThrAsnGlySerIleGlyHisSerProLeuSerLeuSerAlaGl 436
Db 1489 CACCAAGATGCCCAATACCAATGGAAGTATTGGCCACACAGTCCACTTTCTCTGTCAGCCA 1548
Qy 436 nSerValMetGluGluLeuAsnThrAlaProValGlnGluSerProProLeuAlaMetPr 456
Db 1549 GTCTGTAATGGAAGAGCTAAACACTGCACCCGCTCCAGAGAGTCCACCTTGGCCATGCC 1608
```

```
QY 456 oProGlyAsnSerHisGlyLeuGluValcylSerLeuAlaGluValLysGluAsnProPr 476
DB 1609 TCCTGGGAACCTACATGCTAGAGTGGGCTCATTTGGCTGAAGTTAAGGAAACCTCC 1668
QY 476 oPheTyrGlyValIleArgTTPiLeGlyGlnProProGlyLeuAsnGluValLeuAlaG1 496
DB 1669 TTTCTATGGGTAATCCGTTGGATCGGTCAGCAGCCAGGACTGAATGAAGTCTCGCTGG 1728
QY 496 yLeuGluLeuGluAspGluCysAlaGlyCysThrAspGlyThrPheArgGlyThrArgTy 516
DB 1729 ACTGGAACCTGGAAGATGATGTGCGAGGCTGTACGATGGAACCTTCAGAGGCACTCGGTA 1788
QY 516 rPheThrCysAlaLeuLysLysAlaLeuPheValLysLysSerCysArgProAspSe 536
DB 1789 TTTCACTGTGCCCCGTAAGAGGCGCTGTTGTGAACATGAAGAGCTGCAGGCGCTGACTC 1848
QY 536 rArgPheAlaSerLeuGlnProValSerAsnGlnIleGluArgCysAsnSerLeuAlaPh 556
DB 1849 TAGGTTTGCATCATTTGCGAGCGGTTTCCAAATCAGATTGAGCGCTGTAACTCTTTAGCAT 1908
QY 556 eGlyGlyTyrLeuSerGluValValGluGluAsnThrProProLysMetGluLysGluG1 576
DB 1909 TGGAGGCTACTTAAGTGAAGTAGTAGAAGAAAATACTCCACCAAAAATGGAAGAAAGG 1968
QY 576 yLeuGluIleMetIleGlyLysLysGlyIleGlnGlyHisTyrAsnSerCysTyrIle 596
DB 1969 CTTGAGATAATGATTGGGAAGAAAGAGGCACTCAGGCTCAATTAACAATTTCTGTACTT 2028
QY 596 uAspSerThrLeuPheCysLeuPheAlaPheSerSerValLeuAspThrValLeuLeuAr 616
DB 2029 AGACTCAACCTTATTCGCTTATTTGCTTTTGTAGTTCTGTTCTGGACACTGTTACTTAG 2088
QY 616 gProLysGluLysAsnAspValGluTyrTyrSerGluThrGlnGluLeuLeuArgThrG1 636
DB 2089 ACCCAAGAAAGAACGATGAGATATATTATAGTGAACCCCAAGAGCTACTGAGGACAGA 2148
QY 636 uIleValAsnProLeuArgIleTyrGlyTyrValCysAlaThrLysIleMetLysLeuAr 656
DB 2149 AATTGTTAATCCTCTGAGAATATATGGAATATGTGTGTGCCACAAAATTTATGAAACTGAG 2208
QY 656 gLysIleLeuGluLysValGluAlaAserGlyPheThrSerGluGluLysAspProG1 676
DB 2209 GAAAATCTTGAAAGGTGGAGGCTGCATCAGATTACCTCTGAAGAAAAGATCCTGA 2268
QY 676 uGluPheLeuAsnIleLeuPheHisIleLeuArgValGluProLeuLeuLysIleAr 696
DB 2269 GGAATCTTGAATATCTGTTTCATCATATTTTAAGGGTAGAACCTTTGTCTAAAAATAAG 2328
QY 696 gSerAlaGlyGlnLysValGlnAspCysTyrPheTyrGlnIlePheMetGluLysAsnG1 716
DB 2329 ATCAGCAGGTCAAAAGGTACAGATGTGTTACTTCTATCAAAATTTTATGGAAGAAATGA 2388
QY 716 uLysValGlyValProThrIleGlnGlnLeuLeuGluTyrPheIleAsnSerAsnLe 736
DB 2389 GAAAGTTGGCGTTCACAAATTCACAGTTGTAGAAATGGTCTTTTATCAACAGTAACCT 2448
QY 736 uLysPheAlaGluAlaProSerCysLeuIleIleGlnMetProArgPheGlyLysAspPh 756
DB 2449 GAAATTTGCAGAGGCCACCATCATGCTGATTATTCAGATGCTCGATTGGAAGAGACTT 2508
QY 756 eLysLeuPheLysIleIlePheProSerLeuGluLeuAsnIleThrAspLeuLeuGluAs 776
DB 2509 TAAACTATTTAAAAAATTTTTTCCCTCTCTGGAATTAATAATAACAGATTATCTTGAAGA 2568
QY 776 pThrProArgGlnCysArgIleCysGlyGlyLeuAlaMetTyrGluCysArgGlyCysTy 796
DB 2569 CACTCCCAGACAGTCCCGGATGTGGAGGGCTTCCAATGTATGAGTGTAGAGATGCTA 2628
QY 796 rAspAspProAspIleSerAlaGlyLysIleLysGlnPheCysLysThrCysAsnThrG1 816
DB 2629 CGACGATCCGACATCTCAGCTGGAAANAATCAAGCAGTTTCTTAAACCTCGCAACTCA 2688

816 nValHisLeuHisProLysArgLeuAsnHisLysTyrAsnProValSerLeuProLysAs 836
DB 2689 AGTCCACCTTCATCCGAAGAGGCTGAATCATATAATATATACCCAGTGTCACTTCCCAAGA 2748
QY 836 pLeuProAspThrAspThrArgHisGlyCysIleProCysGlnAsnMetGluLeuPheAl 856
DB 2749 CTTACCCGACTGGGACTGGAGACACGGCTGCATCCCTTCCAGAAATATGGAGTTATTGTC 2808
QY 856 aValLeuCysIleGluThrSerHisTyrValAlaPheValLysTyrGlyLysAspSe 876
DB 2809 TGTCTCTGCATAGAAACAAGCCACTATGTTCTTTTGAAGATATGGGAAGGACCATTC 2868
QY 876 rAlaThrLeuPhePheAspSerMetAlaAspArgAspGlyGlyGlnAsnGlyPheAsnI1 896
DB 2869 TGCTGGGCTCTCTTTTGACAGCATGGCCGATCGGATGGTGTGCAGAAATGGCTTCAACAT 2928
QY 896 eProGlnValThrProCysProGluValGlyGluTyrLeuLysMetSerLeuGluAspLe 916
DB 2929 TCTCAAGTCAACCCATGCCAGAGTAGGAGAGTACTTGAAGATGCTCTGGAAGACCT 2988
QY 916 uHisSerLeuAspSerArgArgIleGlnGlyCysAlaArgArgLeuLeuCysAspLaTy 936
DB 2989 GCATTTCTTGGACTCCAGGAGAAATCCAGGCTGTGCAGAGACTGCTTTGTGATGCATA 3048
QY 936 rMetCysMetTyrGlnSerProThrMetSerLeuTyrLys 949
DB 3049 TATGTGCATGTACCAGAGTCCAAATGAGTTTGTACAAA 3088

RESULT 8
US-09-396-970-7559
; Sequence 7559, Application US/09396970
; GENERAL INFORMATION:
; APPLICANT: Gearing, David P.
; APPLICANT: Kingsbury, Gillian A.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES DERIVED FROM A
; TITLE OF INVENTION: MIXED LYMPHOCTE LIBRARY
; FILE REFERENCE: MLN98-40DA
; CURRENT APPLICATION NUMBER: US/09/396,970
; CURRENT FILING DATE: 1999-09-14
; EARLIER APPLICATION NUMBER: 60/100,293
; EARLIER FILING DATE: 1999-09-14
; NUMBER OF SEQ ID NOS: 8756
; SOFTWARE: FastSeq for Windows, Version 3.0
; SEQ ID NO 7559
; LENGTH: 4668
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(4668)
; OTHER INFORMATION: n = A,T,C or G
US-09-396-970-7559

Alignment Scores:
Pred. No.: 0 Length: 4668
Score: 754.00 Matches: 948
Percent Similarity: 98.54% Conservative: 0
Best Local Similarity: 98.54% Mismatches: 1
Query Match: 79.45% Indels: 14
DB: 21 Gaps: 0

US-09-671-687A-3 (1-949) x US-09-396-970-7559 (1-4668)
QY 1 MetSerSerGlyLeuTrpSerGlnGluLysValThrSerProTyrTrpGluGluArgile 20
DB 230 ATGAGTTCAGGCTTATGGAGCCAAAGAAAGTCACTTCCCTTACTGGGAAGCGGATT 289
QY 21 PheTyrLeuLeuLeuGlnGluCysSerValThrAspLysGlnThrGlnLysLeuLeuLys 40
DB 290 TTTTACTTCTCTTCAAGATGCGAGGTTACAGACAAACAAACAAAGCTCTCTTAA 349
QY 41 ValProLysGlySerIleGlyGlnTyrIleGlnAspArgSerValGlyHisSerArgile 60
```

350 GTACCGAAGGAAGTATAGGACAGTATATCAAGATCGTTCTGTGGGCGCATTCAGAGATT 409  
61 ProSerAlaLysGlyLysLeuAsnGlnIleGlyLeuLysIleLeuGluGlnProHisAla 80  
410 CCTTCGCAAAAGCAAGAAATACAGATTGGATTAAAAATCTAGACGACCTCATGCA 469  
81 ValLeuPheValAspGlu---AspValValGluIleAsnGluLysPheThrGluLeuLeu 99  
470 GTTCTCTTTGTGATGAAGAGATGTTCTAGAGATAAAATGAAAAGTTACACAGAGTTACTT 529  
100 LeuAlaIleThrAsnCysGluGluArgPheSerLeuPheLysAsnAsnArgLeuSer 119  
530 TTGGCAATTAACAATTGTGAGAGAGGTTACGCTGTTTAAAAACAGAAACACACATTAAGT 589  
120 LysGlyLeuGlnIleAspValGlyCysProValLysValGlnLeuArgSerGlyGluGlu 139  
590 AAAGGCTCCAAATAGACGTGGCTGTTCTGTGAAAGTACAGCTGAGATCTGGGGAAGAA 649  
140 LysPheProGlyValValArgPheArgGlyProLeuLeuAlaGluArgThrValSerGly 159  
650 AAATTTCTGGAGTTGTACGCTTCAGAGGACCCCTGTTAGCAGAGGACAGCTCTCCGA 709  
160 IlePhePheGlyValGluLeuLeuGluGluArgGlyArgGlyGlnGlyPheThrAspGlyVal 179  
710 ATATTTCTTGAGTTGAAATTGCTGAAGAAGGTGCTGTGTCAGGTTTCACTGACCGGGTG 769  
180 TyrGlnGlyLysGlnLeuPheGlnCysAspGluAspCysGly---PheValAlaLeuAsp 198  
770 TACCAAGGAACACAGCTTTTTCAGTGTGATGAAAGATTGTGGCGTGTGTTGTCATTGGAC 829  
199 LysLeuGluLeuIleGluAspAspThrAlaLeuGluSerAspTyrAlaGlyProGly 218  
830 AAGCTAGAACTCATAGAGATGATGACACTGCATTGGAAAGTGAATTAACGAGGCTCCTGGG 889  
219 AspThrMetGlnValGluLeuProLeuGluIleAsnSerArgValSerLeuLysGly 238  
890 GACCAATGCGAGGTGCAATCTCTCTCTTGGAAATAAATCTCAGAGTTTCTTTTGAA-GGT 948  
239 -GlyGluThrIleGluSerGlyThrValIlePheCysAspValLeuProGlyLysGluSe 258  
949 TGGAGAAACAATAGAACTCGAACAGTTATATCTGTGATGTTTGGCAGGAAGAAAG 1008  
258 rLeuGlyTyrPheValGlyValAspMetAspAsnProIleGlyAsnTrpAspGlyArgPh 278  
1009 CTTAGGATATTTGTTGGTGGACATGATGATAACCTTATGGCAACTGGGATGGAAGATT 1068  
278 eAspGlyVal---LeuCysSerPheAlaCysValGluSerThrIleLeuLeuHisIleAs 297  
1069 TGATGGAGTGCAGCTTTGTAGTTTGGCGTGTGTTGAAAGTACAAATTCATTGCAATCAA 1128  
297 nAspIleIleProGluSerValThrGlnGluArgProProLysLeuAlaPheMetSe 317  
1129 TGATATCATCCAGAGAGTGTGACGAGGAAGAGGCGCTCCCAACTTGCTTTATGTC 1188  
317 rArgGlyValGlyAspLysGlySerSerHisAsnLysProLysAlaThrGlySerTh 337  
1189 AAGAGGTGTTGGGACAAAAGTTTCATCCAGCATATAATAACCAAGGCTACAGGATCTAC 1248  
337 rSerAspProGlyAsnArg---ArgSerGluLeuPheTyrThrLeuAsnGlySerSerVa 356  
1249 CTCAGACCTCGGAATAGAAACAGATCTGAAATATTTTATACCTTAAATGGGTCTCTGT 1308  
356 lAspSerGlnProGlnSerLysAsnThrTrpTyrIleAspGluValAlaGluAs 376  
1309 TGACTCACACCAACCAATCCAATCAAAATAATACATGGTACATGATGAAGTTGCAAGAGA 1368  
376 pProAlaLysSerLeuThrGluIleSerThrAspPheAspArgSerSerProProLeuG1 396  
1369 CCCTGCAAAATCTCTTACAGAGATATCTACAGACTTTGACCGTCTTCCACCACTCCA 1428  
396 nProProProValAsnSerLeuThrThrGluAsnArgPheHisSerLeuProPheSerLe 416  
1429 GCCTCTCTCTGTGAACCTCACTGACCGGAGAACAGATTCCACTCTTTACCATTCAGTCT 1488

416 uThrLysMetProAsnThrAsnGlySerIleGlyHisSerProLeuSerLeuSerAlaG1 436  
1489 CACCAAGATGCCCAATACCAATGGAAGTATTGGCCACACAGTCCACTTTCTCTGTGACGCCA 1548  
436 nSerValMetGluGluLeuAsnThrAlaProValGlnGluSerProProLeuAlaMetPr 456  
1549 GTCTGTAAATGGAAGAGCTAAACACTGCAACCGCTCCAGAGAGTCCACCCCTGGCCATGCC 1608  
456 oProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLysGluAsnProPr 476  
1609 TCCTGGGAACATCACATGGTCTAGAAAGTGGGCTCATTTGGCTGAAGTTAAGGAGAAACCCCTCC 1668  
476 oPheTyrGlyValIleArgTrpIleGlyGlnProProGlyLeuAsnGluValLeuAlaG1 496  
1669 TTTCTATGGGGTATCCGTTGGATCGGTACGCCACCAGGACTGTAATGAAGTGTCTCGCTGG 1728  
496 yLeuGluLeuGluAspGluCysAlaGlyCysThrAspGlyThrPheArgGlyThrArgTy 516  
1729 ACTGNACTGGAAGATGAGTGTGCAAGCTGTACGGATGGAACCTTCAGAGGCACTCGGTA 1788  
516 rPheThrCysAlaLeuLysLysAlaLeuPheValLysLeuLysSerCysArgProAspSe 536  
1789 TTTCACTGTGCTGAAAGAGGCGCTGTTGTGAAACTGAAGAGCTCGAGGCTGACTC 1848  
536 rArgPheAlaSerLeuGlnProValSerAsnGlnIleGluArgCysAsnSerLeuAlaPh 556  
1849 TAGGTTTCATCATTCAGCCGGTTTCCATCAGATTGAGCGCTGTAATCTTTAGCATT 1908  
556 eGlyGlyTyrLeuSerGluValValGluGluAsnThrProProLysMetGluLysGluG1 576  
1909 TGGAGGCTACTTAAGTGAAGTAGTAGAAGAAATACTCCACAAAATGGAAGAAAGG 1968  
576 yLeuGluIleMetIleGlyLysLysGlyIleGlnGlyHisTyrAsnSerCysTyrLe 596  
1969 CTTGAGATAATGATTGGGAAGAAAGGAGCATCCAGGCTATTACAAATTTCTTTGTTACTT 2028  
596 uAspSerThrLeuPheCysLeuPheAlaPheSerSerValLeuAspThrValLeuLeuAr 616  
2029 AGACTCAACCTTATTCCTCTTATTTGCTTTTAGTTCTGTTCTGGACACTGTGTTACTTAG 2088  
616 gProLysGluLysAsnAspValGluTyrTyrSerGluThrGlnGluLeuLeuArgThrG1 636  
2089 ACCCAAGAAAGAACGATGTAGATAATTATAGTGAACCCCAAGAGCTACTGAGGACAGA 2148  
636 uIleValAsnProLeuArgIleTyrGlyTyrValCysAlaThrLysIleMetLysLeuAr 656  
2149 AATTGTTAATCCTCTGAGAATATATGGATATGTGTGTGTCACAAAATAATATGAAACTGAG 2208  
656 gLysIleLeuGluLysValGluAlaAlaSerGlyPheThrSerGluGluLysAspProG1 676  
2209 GAAATACTTGAAAGGTGGAGGCTGCATCAGATTATACCTCTGAAGAAAAGATCCTGA 2268  
676 uGluPheLeuAsnIleLeuPheHisIleLeuArgValGluProLeuLeuLysIleAr 696  
2269 GGAATCTTGTAATATCTGTTTCATCATATTTTAAGGCTAGAACCTTTGCTAAATAAAG 2328  
696 gSerAlaGlyGlnLysValGlnAspCysTyrPheTyrGlnIlePheMetGluLysAsnG1 716  
2329 ATCAGCAGGTCAAAAGGTACAAGATTGTTACTTCTATCAAAATTTTATGGAAGAAAATGA 2388  
716 uLysValGlyValProThrIleGlnLeuLeuGluTrpSerPheIleAsnSerAsnLe 736  
2389 GAAAGTTGGCGTTCACCAATTCAGCAGTGTGTAGAATGGTCTTTTATCAACAGTAACCT 2448  
736 uLysPheAlaGluAlaProSerCysLeuIleIleGlnMetProArgPheGlyLysAspPh 756  
2449 GAAATTTGCAGAGGCAACCATCATGCTGATTATTCAGATGCTCGATTTGGAAGACTT 2508  
756 eLysLeuPheLysIlePheProSerLeuGluLeuAsnIleThrAspLeuLeuGluAs 776  
2509 TAAACTATTTAAAAAATTTTCTCTCTGGAATTTAAATATAACAGATTTCCTTTGAAGA 2568

Qy 776 pThrProArgGlnCysArgIleCysGlyLeuAlaMetTyrGluCysArgGluCysTy 796  
Db 2569 CACTCCAGACAGTGCAGGATATGAGAGGCTTGCAATGTATGAGTGTAGAGAAATGCTA 2628  
Qy 796 rAspAspProAspIleSerAlaGlyIleLysGlnPheCysLysThrCysAsnThrGl 816  
Db 2629 CGACATCCGGACATCTCAGCTGGAAAAATCAAGCAGTGTGTAACACCTGCAACTCA 2688  
Qy 816 nValHisLeuHisProLysArgLeuAsnHisLysTyrAsnProValSerLeuProLysAs 836  
Db 2689 AGTCACCTTCATCCGAGAGGCTGNAATCAATAATATACCCAGTGTCACTCCCAAGA 2748  
Qy 836 pLeuProAspTrpAspTrpArgHisGlyCysIleProCysGlnAsnMetGluLeuPheAl 856  
Db 2749 CTTACCCGACTGGGACTGGAGACAGGCTGCATCCCTGCCAGAAATGAGGATTTTTCG 2808  
Qy 856 aValLeuCysIleGluThrSerHisTyrValAlaPheValLysTyrGlyLysAspAspSe 876  
Db 2809 TGTTCCTCTGCAATAGAAACAAGCCACTATGTTGCTTTTGTGAAGTATGGGAAGGACGATTC 2868  
Qy 876 rAlaTrpLeuPhePheAspSerMetAlaAspArgAspGlyGlyGlnAsnGlyPheAsnIl 896  
Db 2869 TGCCTGGCTCTCTTTTGACAGCATGGCCGATCGGAGTGTGTCAGAAATGGCTTCAACAT 2928  
Qy 896 eProGlnValThrProCysProGluValGlyGluTyrLeuLysMetSerLeuGluAspLe 916  
Db 2929 TCCTCAAGTCACCCCATGCCAGAGTAGGAGATGTTGAAGATGCTCTGGAAGACCT 2988  
Qy 916 uHisSerLeuAspSerArgArgIleGlnGlyCysAlaArgArgLeuLeuCysAspAlaTy 936  
Db 2989 GCATTCCTTGGACTCCAGAGAAATCCAGGCTGTGCACCAAGACTGCTTTGTGATGCATA 3048  
Qy 936 rMetCysMetTyrGlnSerProThrMetSerLeuTyrLys 949  
Db 3049 TATGTGCATGTACCAGAGTCCAAACAATGAGTTGTACAAA 3088

RESULT 9

US-09-397-424-5019  
; Sequence 5019, Application US/09397424  
; GENERAL INFORMATION:  
; APPLICANT: Gearing, David P.  
; APPLICANT: Pan, Yang  
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES DERIVED FROM A  
; TITLE OF INVENTION: HUMAN OSTEOBLAST LIBRARY  
; FILE REFERENCE: MLN98-45PM  
; CURRENT APPLICATION NUMBER: US/09/397,424  
; PRIOR FILING DATE: 1999-09-15  
; PRIOR APPLICATION NUMBER: 60/100,469  
; PRIOR FILING DATE: 1998-09-15  
; PRIOR APPLICATION NUMBER: 60/106,454  
; PRIOR FILING DATE: 1998-10-30  
; PRIOR APPLICATION NUMBER: 60/107,252  
; PRIOR FILING DATE: 1998-11-05  
; PRIOR APPLICATION NUMBER: 60/132,100  
; PRIOR FILING DATE: 1999-04-30  
; NUMBER OF SEQ ID NOS: 5379  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 5019  
; LENGTH: 4668  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(4668)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-397-424-5019

Alignment Scores:  
Pred. No.: 0 Length: 4668  
Score: 754.00 Matches: 948  
Percent Similarity: 98.54% Conservative: 0  
Best Local Similarity: 98.54% Mismatches: 1  
Query Match: 79.45% Indels: 14

DB: 21 Gaps: 0  
US-09-671-687A-3 (1-949) x US-09-397-424-5019 (1-4668)  
Qy 1 MetSerSerGlyLeuTrpSerGlnGluLysValThrSerProTyrTrpGluGluArgIle 20  
Db 230 ATGAGTTTCAAGCTTATGGAGCCCAAGAAAAAGTCACTTCACCTTCTGGGAAGAGCGATT 289  
Qy 21 PheTyrLeuLeuLeuGlnGluCysSerValThrAspLysGlnThrGlnLysLeuLys 40  
Db 290 TTTTACTTCTCTTCAAGATGTCAGCGTTTACAGACAAACAAACAAAGCTCTTTAAA 349  
Qy 41 ValProLysGlySerIleGlyGlnTyrIleGlnAspArgSerValGlyHisSerArgIle 60  
Db 350 GTACCCAGGGAAGTATAGCAGATATTTCAAGATCGTTCTGTGGGGCATTTCAAGGATT 409  
Qy 61 ProSerAlaLysGlyLysLysAsnGlnIleGlyLeuLysIleLeuGlnProHisAla 80  
Db 410 CCTTCTGCAAAAAGGCCAAGAAAAATCAGATTGGGATTTAAAAATTTCTAGAGCAACCTCATGCA 469  
Qy 81 ValLeuPheValAspGlu--AspValValGluIleAsnGlnLysPheThrGluLeuLeu 99  
Db 470 GTTCTCTTTTGTGATGAAAGAGATGTTGTAGATTAATGAAAAAGTTTCAGAGTTACTT 529  
Qy 100 LeuAlaIleThrAsnCysGluGluArgPheSerLeuPheLysAsnArgAsnArgLeuSer 119  
Db 530 TTGGCAATTTACCAATTTGTGAGGAGAGTTTCAGCTCTGTTTAAAAAACAGAAACAGACTAAAT 589  
Qy 120 LysGlyLeuGlnIleAspValGlyCysProValLysValGlnLeuArgSerGlyGluGlu 139  
Db 590 AAAAGGCTCCAAATAGACGTGGGCTGTCTCTGTGNAAGTACAGCTGAGATCTGGGGAAGAA 649  
Qy 140 LysPheProGlyValValArgPheArgGlyProLeuLeuAlaGluArgThrValSerGly 159  
Db 650 AAAATTTCTGGAGTTTGTACGCTTCAGAGGACCCCTGTTAGCAGAGAGGACAGTCTCCGGA 709  
Qy 160 IlePhePheGlyValGluLeuLeuGluGluGlyArgGlyGlnGlyPheThrAspGlyVal 179  
Db 710 ATATTTCTTTGGAGTTGAATTTGCTGGAAGAGGTCGTGGTCAAGGTTTCACTACGGGGTG 769  
Qy 180 TyrGlnGlyLysGlnLeuPheGlnCysAspGluAspCysGly--PheValAlaLeuAsp 198  
Db 770 TACCAAGGGAACACAGCTTTTCACTGTGATGAAGATTGGCGTGTGTTGTGATTCGAC 829  
Qy 199 LysLeuGluLeuIleGluAspAspAspThrAlaLeuGluSerAspTyrAlaGlyProGly 218  
Db 830 AAGCTAGAACTCATAGAAGATGATGACACTGCAATTCGAAAAGTATTCAGCAGGTCCTGGG 889  
Qy 219 AspThrMetGlnValGluLeuProProLeuGluIleAsnSerArgValSerLeuLysGly 238  
Db 890 GACACAAATGCAGTCCGAACTTCTCTCTTTGGAAATTAACCTCCAGAGTTTCTTTGAA--GGT 948  
Qy 239 -GlyGluThrIleGluSerGlyThrValIlePheCysAspValLeuProGlyLysGluSe 258  
Db 949 TGGAGAAACAATAGATCTGGAACAGTTATTTCTGTGATGTTTTCAGGAGAAAGAAAG 1008  
Qy 258 rLeuGlyTyrPheValGlyValAspMetAspAsnProLleGlyAsnTrpAspGlyArgPh 278  
Db 1009 CTTAGGATATTTTGTGTTGGTGTGACATGGATAACCTTAATTCGCAAGTTCGGAAGATT 1068  
Qy 278 eAspGlyVal---LeuCysSerPheAlaCysValGluSerThrIleLeuLeuHisAla 297  
Db 1069 TGATGGAGTGCAGCTTTGTAGTTTTCGGTGTGTGAAGTACAAATTTCTATTCACATCAA 1128  
Qy 297 nAspIleIleProGluSerValThrGlnGluArgArgProProLysLeuAlaPheMetSe 317  
Db 1129 TGNATATCATCCAGAGAGTGTGACGCGAGGAAGAGGCGCTCCCAAACTTTCCTTTATGTC 1188  
Qy 317 rArgGlyValGlyAspLysGlySerSerHisAsnLysProLysAlaThrGlySerTh 337  
Db 1189 AAGAGGTGTTGGGACAAAGGTTTCATCCAGTCAATAAAACCAAGGCTACAGGATCTAC 1248  
Qy 337 rSerAspProGlyAsnArg---ArgSerGluLeuPheTyrThrLeuAsnGlySerSerVa 356



; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: misc feature  
 ; LOCATION: (1) - (4668)  
 ; OTHER INFORMATION: n = A,T,C or G  
 US-09-397-424A-5019

## Alignment Scores:

Pred. No.: 0 Length: 4668  
 Score: 754.00 Matches: 948  
 Percent Similarity: 98.54% Conservative: 0  
 Best Local Similarity: 98.54% Mismatches: 1  
 Query Match: 79.45% Indels: 14  
 DB: 21 Gaps: 0

US-09-671-687A-3 (1-949) x US-09-397-424A-5019 (1-4668)

Qy 1 MetSerSerGlyLeuTrpSerGlnGluLysValThrSerProTyrTrpGluGluArgile 20  
 Db 230 ATGAGTTCAGGCTTATGGAGCAAGAAAAGTCACCTTCACCTTACTGGGAAGAGCGGATT 289  
 Qy 21 PheTyrLeuLeuLeuGlnCysSerValThrAspLysGlnThrGlnLysLeuLeuLys 40  
 Db 290 TTTTACTTGTCTCTTCAAGATGCAGCTTACAGACAAACAAACACAAAAGCTCTTTAAA 349  
 Qy 41 ValProLysGlySerIleGlyGlnTyrIleGlnAspArgSerValGlyHisSerArgile 60  
 Db 350 GTACCGAAGGGAAGATAGGACAGATATATTCAGATCGTTCGTGGGGCAATTCAGGATT 409  
 Qy 61 ProSerAlaLysGlyLysLysAsnGlnIleGlyLeuLysIleLeuGlnProHisAla 80  
 Db 410 CCTTCTGCCAAAGGCAAGAAAATCAGATTGGATTAAATAATCTAGAGCAACCTCATGCA 469  
 Qy 81 ValLeuPheValAspGlu--AspValValGluIleAsnGlnLysPheThrGluLeuLeu 99  
 Db 470 GTTCTCTTGTGATCAAAAGGATGTTGTAGAGATAAATGAAAAGTTTCACAGAGTTACTT 529  
 Qy 100 LeuAlaIleThrAsnCysGluGluArgPheSerLeuPheIysAsnArgAsnArgLeuSer 119  
 Db 530 TTGGCAATTACCAATTGTGAGAGAGGTTTCAGCCTGTTTAAACACAGAACACAGATAAGT 589  
 Qy 120 LysGlyLeuGlnIleAspValGlyCysProValLysValGlnLeuArgSerGlyGluGlu 139  
 Db 590 AAAGCCCTCCAAATAGACGTGGCTGTCTGTGAAAGTACAGCTGAGATCTGGGGAAGAA 649  
 Qy 140 LysPheProGlyValValArgPheArgGlyProLeuLeuAlaGluArgThrValSerGly 159  
 Db 650 AAATTTCTGTGGAGTTGTACGCTTCAGAGGACCCCTGTTAGCAGAGAGGACAGCTCTCCGA 709  
 Qy 160 IlePhePheGlyValGluLeuLeuGluGluGlyArgGlyGlnGlyPheThrAspGlyVal 179  
 Db 710 ATATTCTTTGGAGTTGAATTGCTGGAAGAGGTCGTGGTCAAGGTTTCACTGACGGGGTG 769  
 Qy 180 TyrGlnGlyLysGlnLeuPheGlnCysAspGluAspCysGly--PheValAlaLeuAsp 198  
 Db 770 TACCAGGGAACAGCTTTTTCAGTGTGATGAAGATTGGCGCTGTTGTTCATTGGAC 829  
 Qy 199 LysLeuGluLeuIleGluAspAspThrAlaLeuGluSerAspTyrAlaGlyProGly 218  
 Db 830 AAGCTAGAACTCATAGAAGATGATGACATGCTGATTGGAAAGTGATTACCGAGGCTCTGGG 889  
 Qy 219 AspThrMetGlnValGluLeuProProLeuGluIleAsnSerArgValSerLeuLysGly 238  
 Db 890 GACACAAATGCAGGTCGAATCTCTCTCTTGGAAATAAATCTCCAGAGTTTCTTTGAA--GGT 948  
 Qy 239 GlyGlnThrIleGluSerGlyThrValIlePheCysAspValLeuProGlyLysGluSe 258  
 Db 949 TGGAGAAAACAATAGAAATCTGGAAACAGTTATATCTCTGATGTTTTCGCCAGGAAGAAAG 1008  
 Qy 258 rleuGlyTyrPheValGlyValAspMetAspAsnProIleGlyAsnTrpAspGlyArgPh 278  
 Db 1009 CTTAGGATATTTTGTGTGGACATGATGATTAACCTATTTGGCACTTGGATGGAAGATT 1068

Qy 278 eAspGlyVal---LeuCysSerPheAlaCysValGluSerThrIleLeuLeuHisIleAs 297  
 Db 1069 TGATGGAGTGCAGCTTTGTAGTTTGGTGTGTTGAAAGTACAAATCTATTTCACATCAA 1128  
 Qy 297 nAspIleIleProGluSerValThrGlnGluArgArgProProLysLeuAlaPheMetSe 317  
 Db 1129 TGATATCATCCAGAGAGTGTGACGAGGAAGAGGCGCTCCCAAACTTGCTTTATGTC 1188  
 Qy 317 rArgGlyValGlyAspLysGlySerSerHisAsnLysProLysAlaThrGlySerTh 337  
 Db 1189 AAGAGTGTGTGGGACAAAGGTTTCATCCAGTCTATATATAACCAAGGCTACAGGATCTAC 1248  
 Qy 337 rSerAspProGlyAsnArg---ArgSerGluLeuPheTyrThrLeuAsnGlySerSerVa 356  
 Db 1249 CTCAGACCCCTGGAAATAGAAACAGATCTGAATATTTTATACCTTAATGGGTCTTCTGT 1308  
 Qy 356 lAspSerGlnProGlnSerLysSerLysAsnThrTrpTyrIleAspGluValAlaGluAs 376  
 Db 1309 TGACTCACACCAACAATCCAAATCAAAAATACATGGTACATTCATGAAGTTGCAGAAGA 1368  
 Qy 376 pProAlaLysSerLeuThrCluileSerThrAspPheAspArgSerSerProProLeuGl 396  
 Db 1369 CCCTGCAAAATCTCTTACAGAGATATCTACAGACTTTTGACCGTTCCTTCCACCACCTCCA 1428  
 Qy 396 nProProProValAsnSerLeuThrThrGluAsnArgPheHisSerLeuProPheSerLe 416  
 Db 1429 GCCTCTCTCTGTGTAATCTACTGACCAACGAGACAGATTCACCTCTTTCATTCATTCAGTCT 1488  
 Qy 416 uThrLysMetProAsnThrAsnGlySerIleGlyHisSerProLeuSerLeuSerAlaGl 436  
 Db 1489 CACCAAGATGCCCAATACCAATGGAAGTATTGGCCACAGTCCACTTCTCTGTCTGAGGCCA 1548  
 Qy 436 nSerValMetGluGluLeuAsnThrAlaProValGlnGlnGluSerProProLeuAlaMetPr 456  
 Db 1549 GTCTGTAAATGGAAAGAGCTAAACATGCACCCGTCGAAGAGAGTCCACCTTGGCCATGCC 1608  
 Qy 456 oProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLysGluAsnProPr 476  
 Db 1609 TCTGGAACTCATGTCTAGAGTGGGCTCATTTGGCTGAAAGTTAAGAGAGAACCTTCC 1668  
 Qy 476 oPheTyrGlyValIleArgTrpIleGlyGlnProProGlyLeuAsnGluValLeuAlaGl 496  
 Db 1669 TTTCTATGGGTAAATCCGTTGGATCGGTACGCCACCCAGGACTGAATGAAGTCTCGCTGG 1728  
 Qy 496 yLeuGluLeuGluAspGluCysAlaGlyCysThrAspGlyThrPheArgGlyThrArgTy 516  
 Db 1729 ACTGGAACCTGGAAGATGAGTGTGAGGCTGTACGGATGGAACCTTCAGAGGCACTCGGTA 1788  
 Qy 516 rPheThrCysAlaLeuLysLysAlaLeuPheValLysLeuLysSerCysArgProAspSe 536  
 Db 1789 TTTTCACCTGTGCCCTGAAGAAGCGCTGTTTGTGAAACTGAAGAGCTGCAAGGCTTGACTC 1848  
 Qy 536 rArgPheAlaSerLeuGlnProValSerAsnGlnIleGluArgCysAsnSerLeuAlaPh 556  
 Db 1849 TAGTTTTCATCATTCAGCGCGGTTTCCATCATGATTGAGCGCTGTAACTCTTTAGCATT 1908  
 Qy 556 eGlyGlyTyrLeuSerGluValValGluGluAsnThrProProLysMetGluLysGluGl 576  
 Db 1909 TGGAGGCTACTTAAGTGAAGTAGTAGAAGAAAATACTCCACCAAAAATGGAAGAAAGG 1968  
 Qy 576 yLeuGluIleMetIleGlyLysLysGlyIleGlnGlyHisTyrAsnSerCysTyrLe 596  
 Db 1969 CTTGGAGATAATGATTGGGAAGAAAGAGCATCCAGGGGTCATTACAAATTTCTTGTACTT 2028  
 Qy 596 uAspSerThrLeuPheCysLeuPheAlaPheSerSerValLeuAspThrValLeuLeuAr 616  
 Db 2029 AGACTCAACCTTATCTGCTTATTTGCTTTTAGTTCTGTTCTGGACACTGTGTACTTAG 2088  
 Qy 616 gProLysGluLysAsnAspValGluTyrTyrSerGluThrGlnGluLeuLeuArgThrGl 636  
 Db 2089 ACCCAAGAAAAGAACAGATGATGATATATATAGTGAACCAACAGAGCTACTGAGACAGA 2148



```
QY 636 uileValAsnProLeuArgIleTyGlyTyrValCysAlaThrLysIleMeLysLeuAr 656
Db 2149 AATTGTTAATCTCTGAGAATATATGGATATGTGTGCCACAAAAATATCAAACTGAG 2208
QY 656 gLysIleLeuGluLysValGluAlaSerGlyPheThrSerGluLysLeuAspProGl 676
Db 2209 GAAATACCTTGAAGAGGTGGAGGCTGCATCAGATTACCTCTGAAGAAAAAGATCCTGA 2268
QY 676 uGluPheLeuAsnIleLeuPheHisIleLeuArgValGluProLeuLeuLysIleAr 696
Db 2269 GGAATTCCTGATATCTGTTTCATCATATTTTAAGGGTAGAACCTTTGCTAAATAAG 2328
QY 696 gSerAlaGlyGlnLysValGlnAspCysTyrPheTyrGlnIlePheMetGluLysAsnGl 716
Db 2329 ATCAGAGGTCAAAAGGTACAGATTGTACTCTATCAAAATTTTATGGAATAAATGA 2388
QY 716 uLysValGlyValProThrIleGlnLeuLeuGluTyrSerPheIleAsnSerAsnLe 736
Db 2389 GAAAGTTGGCGTCCCAAAATTCAGCAGTGTGTAGAAATGGTCTTTTATCAACAGTAACCT 2448
QY 736 uLysPheAlaGluAlaProSerCysLeuIleIleGlnMetProArgPheGlyLysAspPh 756
Db 2449 GAAATTCGAGGAGGCACATCATCTGATTAATCAGATGCCCTCGATTGGAAAGACTT 2508
QY 756 eLysLeuPheLysIlePheProSerLeuGluLeuAsnIleThrAspLeuGluAs 776
Db 2509 TAAACTATTTAAAAAATTTTCTCTCTCGAAATTAATAAACAAGATTATTACTTGAAGA 2568
QY 776 pThrProArgGlnCysArgIleCysGlyGlyLeuAlaMetTyrGluCysArgGluCysTy 796
Db 2569 CACTCCACAGACAGTCCCGGATATGGAGGGCTTGCATATGATGATGATGATGATGATG 2628
QY 796 rAspAspProAspIleSerAlaGlyLysIleLysGlnPheCysLysThrCysAsnThrGl 816
Db 2629 CGACGATCGGCACATCTCAGCTGGAAAATAAAGCAGATTTTGTAAACCTCGCAACACTCA 2688
QY 816 nValHisLeuHisProLysArgLeuAsnHisLysTyrAsnProValSerLeuProLysAs 836
Db 2689 AGTCCACCTTCATCCGAAGAGCTGAATCATATAATATAACCCAGTGTCACTTCCCAAGA 2748
QY 836 pleuProAspThrAspThrArgHisGlyCysIleProCysGlnAsnMetGluLeuPheAl 856
Db 2749 CTTACCCAGCTGGGAGCTGGAGACAGCGCTGCATCCCTTGCAGAAATATGAGTATTGTC 2808
QY 856 aValLeuCysIleGluThrSerHisTyrValAlaPheValLysTyrGlyLysAspSe 876
Db 2809 TGTCTCTGCATAGAAACAGCCACTATGTCTTTGTGAAGTATGGGAAGCAGATTC 2868
QY 876 rAlaTrpLeuPhePheAspSerMetAlaAspArgAspGlyGlyGlnAsnGlyPheAsnI 896
Db 2869 TGCCTGGCTCTTTTGACAGATGGCCGATCGGATGGTGTGATGATGGCTTCAACAT 2928
QY 896 eProGlnValThrProCysProGluValGlyGluTyrLeuLysMetSerLeuGluAspLe 916
Db 2929 TCCTCAAGTCACCCCATGCCAGAGTAGGAGTACTTGAAGATGTCTCTCGGAAGACCT 2988
QY 916 uHisSerLeuAspSerArgArgIleGlnGlyCysAlaArgArgLeuLeuCysAspAlaTy 936
Db 2989 GCATTCTTGGACTCCAGGAGAAATCCAGGCTGTGCACGAAGACTGCTTTGTGATGCATA 3048
QY 936 rMetCysMetTyrGlnSerProThrMetSerLeuTyrLys 949
Db 3049 TATGTGCATGTACCAGAGTCCAAACATGAGTTGTACAAA 3088
```

## RESULT 11

US-09-432-241A-3760

; Sequence 3760, Application US/09432241A

; GENERAL INFORMATION:

; APPLICANT: Gearing, David P.

; APPLICANT: Holtzman, Douglas A.

; TITLE OF INVENTION: NUCLEIC ACID MOLECULES DERIVED FROM A

; TITLE OF INVENTION: HUMAN EPIDERMAL KERATINOCYTE LIBRARY

; FILE REFERENCE: 1600.1004001

```
; CURRENT APPLICATION NUMBER: US/09/432,241A
; CURRENT FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/106,445
; PRIOR FILING DATE: 1998-10-30
; PRIOR APPLICATION NUMBER: 60/107,227
; PRIOR FILING DATE: 1998-11-05
; PRIOR APPLICATION NUMBER: 60/127,182
; PRIOR FILING DATE: 1999-03-30
; PRIOR APPLICATION NUMBER: 60/132,099
; PRIOR FILING DATE: 1999-04-30
; NUMBER OF SEQ ID NOS: 5041
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3760
; LENGTH: 4668
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(4668)
; OTHER INFORMATION: n = A,T,C or G
US-09-432-241A-3760
```

## Alignment Scores:

Pred. No.:	0	Length:	4668
Score:	754.00	Matches:	948
Percent Similarity:	98.54%	Conservative:	0
Best Local Similarity:	98.54%	Mismatches:	1
Query Match:	79.45%	Indels:	14
DB:	22	Gaps:	0

US-09-671-687A-3 (1-949) x US-09-432-241A-3760 (1-4668)

```
QY 1 MetSerSerGlyLeuTrpSerGlnGlyValThrSerProTyrTrpGluGluArgIle 20
Db 230 ATGAGTTCAGGCTTATGAGGCCAAGAAAAGTCATTCCCTACTTGGGAAGCGGATT 289
QY 21 PheTyrLeuLeuGlnGluCysSerValThrAspLysGlnThrGlnLysLeuLys 40
Db 290 TTTTACTTGTCTTCAAGNATGACGCTTACAGACAACAACAACAAGCTCTTAAA 349
QY 41 ValProLysGlySerIleGlyGlnTyrIleGlnAspArgSerValGlyHisSerArgIle 60
Db 350 GTACCGAAGGGAAGTATAGGACAGTATATCAAGATCGTCTCTGTGGGCAATCAAGGATT 409
QY 61 ProSerAlaLysGlyLysLysAsnGlnIleGlyLeuLysIleLeuGluGlnProHisAla 80
Db 410 CTTCTCTGCAAAAGCGCAAGAAAATCAGATTGGATTAAAAATTTCTAGACCACTCATGCA 469
QY 81 ValLeuPheValAspGlu---AspValValGluIleAsnGluLysPheThrGluLeuLeu 99
Db 470 GTTCTCTTTTGTGTAAGAGATGTTGTAGAGATTAATGAAGATTACAGAGTTACTTT 529
QY 100 LeuAlaIleThrAsnCysGluGluArgPheSerLeuPheLysAsnArgLeuSer 119
Db 530 TTGGCAATTAACCAATTTGTGAGGAGGTTACSCCTGTTTAAAAACAGAAACAGACTAAGT 589
QY 120 LysGlyLeuGlnIleAspValGlyCysProValLysValGlnLeuArgSerGlyGluGlu 139
Db 590 AAGAGCTCCAAATAGACGTGGGCTGTCTGTGAAAGTACAGCTCAGATCTGGGAAGAA 649
QY 140 LysPheProGlyValValArgPheArgGlyProLeuLeuAlaGluArgThrValSerGly 159
Db 650 AAAATTTCTGGAGTTGTACGCTTCAGAGGACCCCTGTTAGCAGAGAGGACAGCTCTCCGA 709
QY 160 IlePhePheGlyValGluLeuLeuGluGlyArgGlyGlnGlyPheThrAspGlyVal 179
Db 710 ATATTTCTTGGAGTTGAATTTGCTGGAAGAGGTCGTGGTCAAGGTTTCACTACGCGGGTG 769
QY 180 TyrGlnGlyLysGlnLeuPheGlnCysAspGluAspCysGly---PheValAlaLeuAsp 198
Db 770 TACCAAGGGMAACAGCTTTTTCAGTGTGATGAAGATTGTGGCGTGTGTGTTGTCATTGGAC 829
QY 199 LysLeuGluLeuIleGluAspAspThrAlaLeuLeuSerAspTyrAlaGlyProGly 218
```



Db 830 AAGCTAGAACTCATAGAGATGATGACATGCTGATTTGAAAGTGAATTAACAGGTCCTGGG 889  
Qy AspThrMetGlnValGluLeuProLeuGluIleAsnSerArgValSerLeuLysGly 238  
Db 890 GACACATGCGAGTGCACCTCTCTTTGGAAATAAATCCAGAGTTCTTTGAA-GGT 948  
Qy -GlyGluThrIleGluSerGlyThrValIlePheCysAspValLeuProGlyLysGluSe 258  
Db 949 TGGAGAACATAGAACTGGAACAGTTATTTCTGTGATGTTTTCAGAGGAAAGAAAG 1008  
Qy rleuGlyTyrPheValGlyValAspMetAspAsnProIleGlyAsnTrpAspGlyArgPh 278  
Db 1009 CTTAGGATATTTGTTGGTGTGCACATGCAATACCTATTGGCAACTGGGATGGAAGATT 1068  
Qy easpGlyVal---LeuCysSerPheAlaCysValGluSerThrIleLeuLeuHisIleAs 297  
Db 1069 TGATGGAGTGCAGCTTTGTAGTTTTTGCGTGTGTGAAAGTACAAATTTCTATTGCACATCAA 1128  
Qy nAspIleIleProGluSerValThrGlnGluArgArgProProLysLeuAlaPheMetSe 317  
Db 1129 TGATATCATCCAGAGTGTGCACGAGAAAGAGGCCCTCCAAACTTGGCTTTATGTCT 1188  
Qy rArgGlyValGlyAspLysGlySerSerHisAsnLysProLysAlaThrGlySerTh 337  
Db 1189 AAGAGGTGTGGGCAAGGTTTCATCCAGTCAATAAACAAGAGCTACAGGATCTAC 1248  
Qy rSerAspProGlyAsnArg---ArgSerGluLeuPheTyrThrLeuAsnGlySerSerVa 356  
Db 1249 CTCAGACCTCGAAATAGAAACAGATCTCAATATATTTTATACCTTAAATGGGTCTTCTGT 1308  
Qy lAspSerGlnProGlnSerLysSerLysAsnThrTrpTyrIleAspGluValAlaGluAs 376  
Db 1309 TGACTCAACAACCAATCCAAATCAAAATAATCATGTTACATGATGAAGTTGCAGAGA 1368  
Qy pProAlaLysSerLeuThrGluIleSerThrAspPheAspArgSerSerProLeuGlu 396  
Db 1369 CCTCGCAAAATCTCTACAGAGATATCTACAGACTTTGACCGTTCTTCCACCACCTCCA 1428  
Qy nProProValAsnSerLeuThrThrGluAsnArgPheHisSerLeuProPheSerIle 416  
Db 1429 GCCTCTCTGTGAATCACTACACCGAGAACAGATTCACACTCTTTACCATTCAGTCT 1488  
Qy uThrLysMetProAsnThrAsnGlySerIleGlyHisSerProLeuSerLeuSerAlaG 436  
Db 1489 CACCAGAGATGCCAAATACCAATGGAGATATTGGCCACAGTCCACTTCTGTGCAGCCCA 1548  
Qy nSerValMetGluGluLeuAsnThrAlaProValGlnGluSerProProLeuAlaMetPr 456  
Db 1549 GTCTGTATGGAGAGCTAAACACTGCACCGCTCCAGAGAGTCCACCTTTGGCCATGCC 1608  
Qy oProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLysGluAsnProPr 476  
Db 1609 TCTCGGAACTCACATGCTGTAGAGTGGCTCATTTGGCTGAAGTTTAAAGAGAACCTCC 1668  
Qy oPheTyrGlyValIleArgTrpIleGlyGlnProProGlyLeuAsnGluValLeuAlaG 496  
Db 1669 TTTCTATGGGGTAATCCGTTGATCGGTGAGCCAGCCAGGACTGAATGAAGTGTCTGCTGG 1728  
Qy yLeuGluLeuGluAspGluCysAlaGlyCysThrAspGlyThrPheArgGlyThrArgTy 516  
Db 1729 ACTGGAACCTGGAAGATGAGTGTGAGGCTGTACGGATGGAACCTTCAGAGGCACTCGGTA 1788  
Qy rPheThrCysAlaLeuLysLysAlaLeuPheValLysLeuLysSerCysArgProAspSe 536  
Db 1789 TTTCACTGTGCTCGCTGAAGAGCGCTGTTTGTGAACTGAAGAGCTGCAGGCGCTGACTC 1848  
Qy rArgPheAlaSerLeuGlnProValSerAsnGlnIleGluArgCysAsnSerLeuAlaPh 556  
Db 1849 TAGGTTTGCATCATTTGCAGCGGCTTTCCAATCAGATTGAGCGCTGTAACTCTTTAGCAT 1908  
Qy eGlyGlyTyrLeuSerGluValValGluGluAsnThrProProLysMetGluLysGlu 576

Db 1909 TGGAGGCTACTTAAGTGAAGTAGTAGAAGAAATACTCCACCACCAAAATGGAAGAAAGAGG 1968  
Qy yLeuGluIleMetIleGlyLysLysGlyIleGlnGlyHisTyrAsnSerCysTyrIle 596  
Db 1969 CTTGGAGATAATGATTGGGAAGAAAGAGCATCCAGGGTCAATTACAATCTCTTTACTTT 2028  
Qy uAspSerThrLeuPheCysLeuPheAlaPheSerSerValLeuAspThrValLeuLeuAr 616  
Db 2029 AGACTCAACTTATTCTGCTTATTCTGCTTTAGTTTCTGTTCTGAGACACTGTGTACTTAG 2088  
Qy gProLysGluLysAsnAspValGluTyrTyrSerGluThrGlnGluLeuLeuArgThrGl 636  
Db 2089 ACCCAAGAAAGAACGATGTAGAATATTATAGTGAACCCCAAGAGCTACTGAGGACAGA 2148  
Qy uIleValAsnProLeuArgIleTyrGlyTyrValCysAlaThrLysIleMetLysLeuAr 656  
Db 2149 AATTGTTAATCTCTGAGAAATATATGATATGTGTGTGCCACAAATAATATGAACTGAG 2208  
Qy glyIleLeuGluLysValGluAlaAlaSerGlyPheThrSerGluGluLysAspProGl 676  
Db 2209 GAAATACTTGAAGAGTGGAGCTGCATCAGGATTTACCTCTGAAGAAAGAAAGATCTGA 2268  
Qy uGluPheLeuAsnIleLeuPheHisIleLeuArgValGluProLeuLeuLysIleAr 696  
Db 2269 GGAATTTCTTGAATATTTCTGTTTCATCATATTTTAAAGGTAGAACCTTTGCTAAAAATAAG 2328  
Qy gSerAlaGlyGlnLysValGlnAspCysTyrPheTyrGlnIlePheMetGluLysAsnGl 716  
Db 2329 ATCAGAGGTCAAAAGGTACAGATTGTACTTCTATCAAAATTTTATGGAAAAAATGA 2388  
Qy uLysValGlyValProThrIleGlnLeuLeuGluTrpSerPheIleAsnSerAsnLe 736  
Db 2389 GAAAGTTGGCGTTCCCACTTCAGCAGTTGTTAGAAATGCTCTTTTATCAACAGTAACCT 2448  
Qy uLysPheAlaGluAlaProSerCysLeuIleIleGlnMetProArgPheGlyLysAspPh 756  
Db 2449 GAAATTTGCAGAGGACCATCATGCTGTGATTTATTCAGATGCCCTCGATTTTGGAAAAAGACT 2508  
Qy elysLeuPheLysLysIlePheProSerLeuGluLeuAsnIleThrAspLeuLeuGluAs 776  
Db 2509 TAAACTATTAAAAAAATTTTCTCTCTCGGAATTAATAATATACAGATTTACTTTGAAGA 2568  
Qy pThrProArgGlnCysArgIleCysGlyGlyLeuAlaMetTyrGluCysArgGluCysTy 796  
Db 2569 CACTCCCAAGACGTGCGGATATGTGAGGGCTTGCAATGATGATGATGATGATGATGATG 2628  
Qy rAspAspProAspIleSerAlaGlyLysIleLysGlnPheCysLysThrCysAsnThrGl 816  
Db 2629 CGACGATCCGGACATCTCAGCTGGAATAATCAAGCAGTTTGTAAAAACCTGCAACACTCA 2688  
Qy nValHisLeuHisProLysArgLeuAsnHisLysTyrAsnProValSerLeuProLysAs 836  
Db 2689 AGTCCACCTTCATCCGAAGAGGCTGAATCATATAATATAACCCAGTGTCACTTCCCAAGA 2748  
Qy pLeuProAspTrpAspTrpArgHisGlyCysIleProCysGlnAsnMetGluLeuPheAl 856  
Db 2749 CTTACCCGACTGGAGCTGGAGACACCGCTGCATCCCTTGCAGAAATATGAGATTTATTC 2808  
Qy aValLeuCysIleGluThrSerHisTyrValAlaPheValLysTyrGlyLysAspSe 876  
Db 2809 TGTTCTCTGCATAGAAACCAAGCCACTATGTTGCTTTTGTGAAGTATGGGAAGGACGATTC 2868  
Qy rAlaTrpLeuPhePheAspSerMetAlaAspArgAspGlyGlyGlnAsnGlyPheAsnIl 896  
Db 2869 TGCTGCTCTCTTTTGACAGCATGGCCGATGGGATGGTGGTGCAGAAATGGCTTCAACAT 2928  
Qy eProGlnValThrProCysProGluValGlyGluTyrLeuLysMetSerLeuGluAspLe 916  
Db 2929 TCTCAAGTCACCCCATGCCAGAGTAGGAGATGACTTGAAGATGTCTCTGGAAGACCT 2988  
Qy uHisSerLeuAspSerArgIleGlnGlyCysAlaArgArgLeuLeuCysAspAlaTy 936  
Db 2989 GCATTTCTTGGACTCCAGGAGAAATCCAGGCTGTGCACGAAGACTGCTTTGTGTGTCATA 3048

Qy 936 rMetCysMetTyrGlnSerProThrMetSerLeuTyrLys 949  
 Db 3049 TATGTGCATGTACAGAGTCCAAACATGAGTTTGTACAAA 3088

## RESULT 12

US-09-434-737-1278  
 ; Sequence 1278, Application US/09434737  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Gearing, David P.  
 ; APPLICANT: Kingsbury, Gillian A.  
 ; TITLE OF INVENTION: NUCLEIC ACID MOLECULES DERIVED FROM A  
 ; TITLE OF INVENTION: HUMAN MIXED LYMPHOCYTE LIBRARY  
 ; FILE REFERENCE: 1600.1067001  
 ; CURRENT APPLICATION NUMBER: US/09/434,737  
 ; CURRENT FILING DATE: 1999-11-05  
 ; EARLIER APPLICATION NUMBER: 60/107,228  
 ; EARLIER FILING DATE: 1998-11-05  
 ; NUMBER OF SEQ ID NOS: 1830  
 ; SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ ID NO 1278  
 ; LENGTH: 4668  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: misc feature  
 ; LOCATION: (1)...(4668)  
 ; OTHER INFORMATION: n = A,T,C or G  
 US-09-434-737-1278

## Alignment Scores:

Pred. No.: 0 Length: 4668  
 Score: 754.00 Matches: 948  
 Percent Similarity: 98.54% Conservatives: 0  
 Best Local Similarity: 98.54% Mismatches: 1  
 Query Match: 79.45% Indels: 14  
 DB: 22 Gaps: 0

US-09-671-687A-3 (1-949) x US-09-434-737-1278 (1-4668)

Qy 1 MetSerSerGlyLeuTyrSerGlnGluLysValThrSerProTyrTrpGluGluArgLe 20  
 Db 230 ATGAGTTTCAGGCTTATGGAGGCAAGAAAAGTCACTTCACCTACTGGAGAGCGGATT 289  
 Qy 21 PheTyrLeuLeuLeuGlnGluCysSerValThrAspLysGlnThrClnLysLeuLeuLys 40  
 Db 290 TTTTACTTGTCTTCTCAAGATGCGAGCGTTACAGACAAACAAACAAAGCTCTCTAAA 349  
 Qy 41 ValProLysGlySerIleGlyClnTyrIleGlnAspArgSerValGlyHisSerArgile 60  
 Db 350 GTACCGAAGGGAAGTATAGGACAGTATATCAAGATCGTTCTGTGGGCAATTCAGGATT 409  
 Qy 61 ProSerAlaLysGlyLysLysAsnGlnIleGlyLeuLysIleLeuGluGlnProHisAla 80  
 Db 410 CCTTCTGCAAAAGGCAAGAAAATCAGATTGGATTAAATAATCTAGAGCAACCTCATGCA 469  
 Qy 81 ValLeuPheValAspGlu---AspValValGluIleAsnGluLysPheThrGluLeuLeu 99  
 Db 470 GTTCTCTTTGTGATGAAGAAGATGTTTAGAGATATAATGAAGAATTCACAGAGTTACTT 529  
 Qy 100 LeuAlaIleThrAsnCysGluGluArgPheSerLeuPheLysAsnArgAsnArgLeuSer 119  
 Db 530 TTGGCAATTACCAATTGTGAGGAGAGGTTACGCTGTTTAAACAGAAACAGATTAAGT 589  
 Qy 120 LysGlyLeuGlnIleAspValGlyCysProValLysValGlnLeuArgSerGlyGluGlu 139  
 Db 590 AAAGGCTCCAAATAGACGTGGCTGCTCTGTGAAGTACAGCTGAGATCTGGGAGAA 649  
 Qy 140 LysPheProGlyValValArgPheArgGlyProLeuLeuAlaGluArgThrValSerGly 159  
 Db 650 AAATTCCTGGAGTTGTACGCTTCAGAGACCCCTGTTAGCAGAGAGGACAGCTCCGGA 709  
 Qy 160 IlePheGlyValGluLeuLeuGluGlyArgGlyGlnGlyPheThrAspGlyVal 179

Db 710 ATATTCTTTGGAGTTGAATTGCTGGAAGAAGTCTGCTCAAGTTTTCATCTACGCGGGTG 769  
 Qy 180 TyrGlnGlyLysGlnLeuPheGlnCysAspGluAspCysGly---PheValAlaLeuAsp 198  
 Db 770 TACCAAGGGAACACAGCTTTTTCAGTGTGATGAAGATTGTGGCGTGTGTTGTCATTGGAC 829  
 Qy 199 LysLeuGluLeuIleGluAspAspThrAlaLeuGluSerAspTyrAlaGlyProGly 218  
 Db 830 AAGCTAGAACTCATAGAGATGATGACACTGCTGATTGGAAGTGAATTACGACGCTCTGG 889  
 Qy 219 AspThrMetGlnValGluLeuProLeuGluIleAsnSerArgValSerLeuLysGly 238  
 Db 890 GACACAATGTCAGTCACTCTCTCTTTGGAATAAACTCCAGAGTTTCTTTGAA-GGT 948  
 Qy 239 -GlyGluThrIleGluSerGlyThrValIlePheCysAspValLeuProGlyLysGlu 258  
 Db 949 TGGAGAAACAATAGAACTCTGGAACAGTATATCTGTGATGTTTTCAGGAGAAAGAAAG 1008  
 Qy 258 rLeuGlyTyrPheValGlyValAspMetAspAsnProIleGlyAsnTrpAspGlyArgPh 278  
 Db 1009 CTTAGATATTTTGTGTGGTGGACATGGATTAACCTATTGGCAACTGGGATGGAAGATT 1068  
 Qy 278 eAspGlyVal---LeuCysSerPheAlaCysValGluSerThrIleLeuLeuHisIle 297  
 Db 1069 TGATGGAGTGACGCTTTGTAGTTTTCGCTGTGTTGAAAGTACAATTTCTATTGCACATCAA 1128  
 Qy 297 nAspIleIleProGluSerValThrGlnGluArgArgProProLysLeuAlaPheMetSe 317  
 Db 1129 TGATATATCTCCAGAGAGTGTGACGAGAAAGAGGCTCCCAACTTGGCTTTATGTC 1188  
 Qy 317 rArgGlyValGlyAspLysGlySerSerHisLeuLysProLysAlaThrGlySerTh 337  
 Db 1189 AAGAGTGTGGGGAACAAAGTTTCATCCAGTCAATATAACCAAGGCTTACAGAGTCTAC 1248  
 Qy 337 rSerAspProGlyAsnArg---ArgSerGluLeuPheTyrThrLeuAsnGlySerSerVa 356  
 Db 1249 CTCAGACCTCGAAATAGAAACAGATCTGAATTTATATACCTTAAATGGGTCTTCTGT 1308  
 Qy 356 lAspSerGlnProGlnSerLysSerIleAsnThrTrpTyrIleAspGluValAlaGlu 376  
 Db 1309 TGACTCACACCAACCAATCCAAATCAAAAATACATGTTACATTTGATGAAGTTGCAGA 1368  
 Qy 376 pProAlaLysSerLeuThrGluIleSerThrAspPheAspArgSerSerProLeuGlu 396  
 Db 1369 CCTTCGAAAATCTCTTACAGAGATATCTACAGACTTTGACCGTCTTCTTACCACCACTCA 1428  
 Qy 396 nProProValAsnSerLeuThrThrGluAsnArgPheHisSerLeuProPheSerLe 416  
 Db 1429 GCCTCTCTGTGAACTCACTGACCAACGAGAACAGATTCCACTCTTTTACCATTCACTCT 1488  
 Qy 416 uThrLysMetProAsnThrAsnGlySerIleGlyHisSerProLeuSerLeuSerAlaG 436  
 Db 1489 CACCAAGATGCCCAATACCAATGGAAGTATTGGCCACAGTCCACTTTCTCTGTGACGCC 1548  
 Qy 436 nSerValMetGluGluLeuAsnThrAlaProValGlnGluSerProProLeuAlaMetPr 456  
 Db 1549 GTCTGTAAATGGAAGAGCTAAACACCTGCCACCTCCAGAGAGTCCACCTTGGCCATGCC 1608  
 Qy 456 oProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLysGluAsnProPr 476  
 Db 1609 TCCTGGAACTCACTCACTGTTAGAGTGGGCTCATTTGGCTGAAGTTAAGAGAACCTTCC 1668  
 Qy 476 oPheTyrGlyValIleArgTrpIleGlyGlnProProGlyLeuAsnGluValLeuAlaG 496  
 Db 1669 TTTCTATGGGTAAATCCGTTGGATCGGTGAGTCCAGCAGCAGACTGAATGAAGTCTCGCTG 1728  
 Qy 496 yLeuGluLeuGluAspGluCysAlaGlyCysThrAspGlyThrPheArgGlyThrArgTy 516  
 Db 1729 ACTGGAACCTGGAAGATGAGTGTGAGGCTGTACGGATGGAACCTTTCAGAGGACCTCGGTA 1788  
 Qy 516 rPheThrCysAlaLeuLysLysAlaLeuPheValLysLeuLysSerCysArgProAspSe 536

1789	TTTCACCTGTGGCCCTGAAGAAGCGCGCTGTTTGTGGAAACTGAAGAGCTGCGAGGCGCTGACTC	1848
Qy	rAtqPheAlaSerLeuGlnProValSerAsnGlnIleGluArgCysAsnSerLeuAlaPh	556
Db	TAGGTTTGGCATCAITTCGAGCGGTTTCCAAATCAGATTCAGCGCTGTAACCTCTTTAGCATT	1908
Qy	eGlyGlyTyrLeuSerGluValValGluGluAsnThrProProLysMetGluLysGluGI	576
Db	1909 TGGAGGCTACTTAAGTGAAGTAGTAGAAGAAAAATACTCCACMAAAATGGAAAAAGAAGG	1968
Qy	YLeuGluIleMetIleGlyIlyIstYsGlyIleGlnGlyHisTyrAsnSerCysTyrLe	596
Db	1969 CTTGGAGATTAATGATTGGGAAGAAGAAGGCATCAGGGTCATTACAAATTCCTTGTTTACTT	2028
Qy	uAspSerThrLeuPheCysLeuPheAlaPheSerSerValLeuAspThrValLeuLeuAr	616
Db	2029 AGACTCAACCTATTCTTGCTATTATTGCTTTTAGTTCTGTTCTGGACACTGTGTTACTTAG	2088
Qy	gProLysGluLysAsnAspValGluTyrTyrSerSerGluThrGlnGluLeuLeuArgThrGI	636
Db	2089 ACCCAAGAAAAAGAACGATGTAGAATATTATAGTGAAACCCCAAGAGCTACTGAGGACAGA	2148
Qy	uIleValAsnProLeuArgIleTyrGlyTyrValCysAlaThrIlyIstMetLysLeuAr	656
Db	2149 AATTGTTAAATCCTCTGAGATATATGATATGTGTGTGCCACAAAAATTTATGAACACTGAG	2208
Qy	gLyIleLeuGluLysValGluAlaIleSerGlyPheThrSerGluGluLysAspProGI	676
Db	2209 GAAATTAATCTGAAAAGGTGGAGGCTGCATCAGGATTTTACCTCTGAGAGAAAAAGATCTCTGA	2268
Qy	uGluPheLeuAsnIleLeuPheHisIleLeuArgValGluProLeuLeuLysIleAr	696
Db	2269 GGAATTCCTTGAATATTCTGTTTCATCATATTTTAAAGGTGAACCTTTGCTAAAAAATAAG	2328
Qy	gSerAlaGlyGlnLysValGlnAspCysTyrPheTyrGlnIlePheMetGluLysAsnGI	716
Db	2329 ATCAGCAGGTCAAAAGGTACAAAGATTGTTACTTCTATCAAAATTTTTATGGAATAAAATGA	2388
Qy	uLyValGlyValProThrIleGlnGlnLeuLeuGluTrpSerPheIleAsnSerAsnLe	736
Db	2389 GAAAGTTGGCGGTTCCCAATTCAGCAGTGTGTAGATGGTGCTTTTATCAACAGTAACCT	2448
Qy	uLyPheAlaGluAlaProSerCysLeuIleIleGlnMetProArgPheGlyLysAspPh	756
Db	2449 GAAATTTCCAGAGGCACCATCATGCTGATTAATTACAGATGCTCGATTTTGAAAAAGACTT	2508
Qy	eLysLeuPheLysIstLysIlePheProSerLeuGluLeuAsnIleThrAspLeuLeuAs	776
Db	2509 TAAACTATTTAAAAAATTTTTTCTCTCTGGAATTAATAATATAACAGATTTTACTTTGAAGA	2568
Qy	pThrProArgGlnCysArgIleCysGlyGlyLeuAlaMetTyrGluCysArgGluCysTy	796
Db	2569 CACTCCAGACAGTGC CGGATATGTGAGGGCTTGCAATGTATGAGTGTATGAGAGTAATGCTA	2628
Qy	rAspAspProAspIleSerAlaGlyLysIleLysGlnPheCysLysThrCysAsnThrGI	816
Db	2629 CGACGATCCGACATCTCAGCTGGAAAAATCAACGAGTTTTTGTAAAAACCTTGCAACACTCA	2688
Qy	nValIleLeuHisProLysArgLeuAsnHisLysTyrAsnProValSerLeuProLysAs	836
Db	2689 AGTCCACCTTCATCCGAAGAGGCTGAATCATATAATAAATATAACCCAGTGTCACTTCCCAAGA	2748
Qy	pLeuProAspTrpAspTrpArgHisGlyCysIleProCysGlnAsnMetGluLeuPheAl	856
Db	2749 CTTTACCAGACTGGGACCTGGAGACACGCGTGCATCCCTTGGCAGAAATATGAGAGTTATTTCG	2808
Qy	aValLeuCysIleGluThrSerHisTyrValAlaPheValLysTyrGlyLysAspAspSe	876
Db	2809 TGTTCCTGCATAGAAACACGCACTATGTGCTTTTGTGAGTATGGGAAGGACGATTC	2868
Qy	rAlaTrpLeuPhePheAspSerMetAlaAspArgAspGIyGlyGlnAsnGlyPheAsnII	896
Db	2869 TGCCTGGCTCTCTTTTGACAGCATGGCGGATCGGGATGGTGGTCAATGGCTTCAACAT	2928

```

Qy 896 eProGlnValThrProCysProGluValGlyGluTyrLeuLysMetSerLeuGluAspLe 916
Db 2929 TCCTCAAGTCACCCCATGCCAGAGTAGGAGAGTACTTTGAAGATGCTCTGGAAGACCT 2988
Qy 916 uHisSerLeuAspSerArgArgIleGlnGlyCysAlaAacArgLeuLeuCysAspAlaTy 936
Db 2989 GCATTCCTCGGACTCCAGGAGATCCAAAGCTGTGCACGAAGACTGCTTTGTGATGCATA 3048
Qy 936 rMetCysMetTyrGlnSerProThrMetSerLeuTyrLys 949
Db 3049 TATGTGCATGTACCAGAGTCCAAATGAGTTGTATCAAA 3088

RESULT 13
US-09-850-118-1278
; Sequence 1278, Application US/09850118
; GENERAL INFORMATION:
; APPLICANT: Gearing, David P.
; APPLICANT: Kingsbury, Gillian A.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES DERIVED FROM A
; TITLE OF INVENTION: HUMAN MIXED LYMPHOCYTE LIBRARY
; FILE REFERENCE: 1600.1067-002
; CURRENT APPLICATION NUMBER: US/09/850,118
; CURRENT FILING DATE: 2001-05-07
; PRIOR APPLICATION NUMBER: US 60/107,228
; PRIOR FILING DATE: 1998-11-05
; PRIOR APPLICATION NUMBER: US 09/434,737
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 1830
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1278
; LENGTH: 4668
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(4668)
; OTHER INFORMATION: n = A,T,C or G
US-09-850-118-1278

```

Alignment Scores:		
Pred. No.:	0	Length: 4668
Score:	754.00	Matches: 948
Percent Similarity:	98.54%	Conservative: 0
Best Local Similarity:	98.54%	Mismatches: 1
Query Match:	79.45%	Indels: 14
DB:	36	Gaps: 0

120 LysGlyLeuGlnIleAspValGlyCysProValLysValGlnLeuArgSerGlyGluGlu 139  
Db  
590 AAAGGCTCCAAATAGACGCTGGCTGCTCTGTGAAAGTACAGCTGAGATCTGGGGAAGA 649  
Qy  
140 LysPheProGlyValValArgPheArgGlyProLeuLeuAlaGluArgThrValSerGly 159  
Db  
650 AAATTTCTGAGTGTGATGCTTACAGAGACCCCTGTGTAGCAGAGACAGCTCTCCGGA 709  
Qy  
160 IlePhePheGlyValGluLeuLeuGluGluArgGlyArgGlyGlnGlyPheThrAspGlyVal 179  
Db  
710 ATATTTCTTGGAGTTGAATGCTGGAAGAGGTGCTGTGTCAAGTTTCACTGACGGGGTG 769  
Qy  
180 TyrGlnGlyLysGlnLeuPheGlnCysAspGluAspCysGly---PheValAlaLeuAsp 198  
Db  
770 TACCAAGGGAACAGCTTTTTCAGTGTGATGAAGATTGTGGCGTGTGTTGTCATTGGAC 829  
Qy  
199 LysLeuGluLeuIleGluAspAspThrAlaLeuGluSerAspTyrAlaGlyProGly 218  
Db  
830 AAGCTAGAACTCATAGAAAGATGATGACACTGCTGCAATGGAAAGTGAATACGCAAGTCTCGG 889  
Qy  
219 AspThrMetGlnValGluLeuProLeuGluIleAsnSerArgValSerLeuLysGly 238  
Db  
890 GACAAATGCAAGTGGAACTTCTCTTGGAAATAAACTCCAGAGTTTCTTTGAA-GGT 948  
Qy  
239 -GlyGluThrIleGluSerGlyThrValIlePheCysAspValLeuProGlyLysGluSe 258  
Db  
949 TGGAGAAACAATAGAACTCGNAACAGTTATTTCTGTGATGTTTGGCCAGGNAAGAAAG 1008  
Qy  
258 rLeuGlyTyrPheValGlyValAspMetAspAsnProIleGlyAsnTrpAspGlyArgPh 278  
Db  
1009 CTTAGSATATTTGTGTGTGGACATGATGATGATGATGATGATGATGATGATGATGATGAT 1068  
Qy  
278 eAspGlyVal---LeuCysSerPheAlaCysValGluSerThrIleLeuLeuHisIleAs 297  
Db  
1069 TGATGGAGTGCAGCTTTTGTAGTTTGTGGTGTGTAAGATGATGATGATGATGATGATGAT 1128  
Qy  
297 nAspIleProGluSerValThrGlnGluArgArgProProLysLeuAlaPheMetSe 317  
Db  
1129 TGATATCATCCAGAGAGTGTGACCGAGGAAGAGGCTCCCAACTTGGCTTTATGTC 1188  
Qy  
317 rArgGlyValGlyAspLysGlySerSerHisAsnLysProLysAlaThrGlySerTh 337  
Db  
1189 AAGAGGTGTGGGACAAAGGTTTCATCCAGTCATAATAAACCAAGGCTACAGGATCTAC 1248  
Qy  
337 rSerAspProGlyAsnArg---ArgSerGluLeuPheTyrThrLeuAsnGlySerSera 356  
Db  
1249 CTCAGACCTCGAAATAGAAACAGATCTGAATTTATTTATACCTTTAAATGGGTCTCTGT 1308  
Qy  
356 lAspSerGlnProGlnSerLysSerLysAsnThrTrpTyrIleAspGluValAlaGluAs 376  
Db  
1309 TGACTCAACACACACANTCAAATCAAAATAATACATGGTACATGATGAAGTTGCAAGA 1368  
Qy  
376 pProAlaLysSerLeuThrGluIleSerThrAspPheAspArgSerSerProProLeuG1 396  
Db  
1369 CCCTGCAAAATCTTTACAGAGATATCTACAGACTTTGACCGTCTTTTACCAACCACTCCA 1428  
Qy  
396 nProProProValAsnSerLeuThrThrGluAsnArgPheHisSerLeuProPheSerIe 416  
Db  
1429 GCCTCTCTGTGAACCTCACTGACCCAGAGAACAGATTTCCACTCTTTACCAATTCAGTCT 1488  
Qy  
416 uThrLysMetProAsnThrAsnGlySerIleGlyHisSerProLeuSerLeuSerAlaG1 436  
Db  
1489 CACCAAGATGCCCAATCAATAGGAAGATTATGGCCACAGTCACATCTTCTGTGACGCCCA 1548  
Qy  
436 nSerValMetGluGluLeuAsnThrAlaProValGlnGluSerProProLeuAlaMetPr 456  
Db  
1549 GTCTGTAATGAAGAGCTTAACACTGACCGCTCCAGAGAGTCCACCCCTTGGCCATGCC 1608  
Qy  
456 oProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLysGluAsnProPr 476  
Db  
1609 TCCTGGGAACCTCACATGCTAGAAAGTGGGCTCATTTGGCTGAAGTTAAGGAGAACCCCTCC 1668

Qy  
476 oPheTyrGlyValIleAArgTrpIleGlyGlnProProGlyLeuAsnGluValLeuAlaG1 496  
Db  
1669 TTTCTATGGGGTAATCCGTTGGATCGGTAGGCTCAGCCACAGGACTGAATGAAGTCTCGCTGG 1728  
Qy  
496 yLeuGluLeuGluAspGluCysAlaGlyCysThrAspGlyThrPheArgGlyThrArgTy 516  
Db  
1729 ACTGGAACCTGGAAGATGAGTGTGCAAGGCTGTACGATGAAACCTTCAGAGGCACTCGGTA 1788  
Qy  
516 rPheThrCysAlaLeuLeuLysLysAlaLeuPheValLysLeuLysSerCysArgProAspSe 536  
Db  
1789 TTTACCTGTGCCCTGAAGAGGCGCTGTTGTGTGAACTGAAGAGCTGCAGGCGCTGACTC 1848  
Qy  
536 rArgPheAlaSerLeuGlnProValSerAsnGlnIleGluArgCysAsnSerLeuAlaPh 556  
Db  
1849 TAGGTTTGCATCATTTGACCGCGTTTCCAATCAGATTGAGCGCTGTAACTCTTTAGCATT 1908  
Qy  
556 eGlyGlyTyrLeuSerGluValValGluGluAsnThrProProLysMetGluLysGluG1 576  
Db  
1909 TGGAGGCTACTTAAGTGAAGTAGTAGAAGAAATACTCCACCAAAATGGAAGAAAGAGG 1968  
Qy  
576 yLeuGluIleMetIleGlyLysLysLysGlyIleGlnGlyHisTyrAsnSerCysTyrLe 596  
Db  
1969 CTTGAGATATGATTGGAGAGAGGCAATCCAGGCTCATTAACAATTTCTTTGTTACTT 2028  
Qy  
596 uAspSerThrLeuPheCysLeuPheAlaPheSerSerValLeuAspThrValLeuLeuAr 616  
Db  
2029 AGACTCAACCTTATTTCTGCTTATTTGCTTTTAGTTCTGTTCTGGACACTGTGTTACTTAG 2088  
Qy  
616 gProLysGluLysAsnAspValGluTyrTyrSerGluThrGlnGluLeuLeuArgThrG1 636  
Db  
2089 ACCCAAGAAAGAAAGACGATGTAGATAATTTAGTAGTGAACCCCAAGAGCTACTCAGGACAGA 2148  
Qy  
636 uIleValAsnProLeuArgIleGlyTyrValCysAlaThrLysIleMetLysLeuAr 656  
Db  
2149 AATTGTTAATCTCTGAGAATATATGATATGTGTGTCACCAAAATTTATGAACTGAG 2208  
Qy  
656 glysIleLeuGluLysValGluAlaAlaSerGlyPheThrSerGluGluLysAspProG1 676  
Db  
2209 GAAATATCTTGAAGAGTGGAGGCTGCATCAGGATTTACCTCTGAAGAAAGATCTCTGA 2268  
Qy  
676 uGluPheLeuAsnIleLeuPheHisIleLeuArgValGluProLeuLeuLysIleAr 696  
Db  
2269 GGAAATCTTGAATATTTCTGTTTTCATCATATTTAAGGGTAGAACCTTTGCTAAAAATAAG 2328  
Qy  
696 gSerAlaGlyGlnLysValGlnAspCysTyrPheTyrGlnIlePheMetGluLysAsnG1 716  
Db  
2329 ATCAGCAGGTCAAAAGGTACAGATTTGTTACTTCTATCAAAATTTTATGGAAGAAATGA 2388  
Qy  
716 uLysValGlyValProThrIleGlnGlnLeuLeuGluTrpSerPheIleAsnSerAsnLe 736  
Db  
2389 GAAAGTTGGCGTTCCCAATTCAGAGTTGTAGAAATGCTCTTTTATCAAGATTAACCT 2448  
Qy  
736 uLysPheAlaGluAlaProSerCysLeuIleIleGlnMetProArgPheGlyLysAspPh 756  
Db  
2449 GAAATTTGCAGAGGACCATCATCTGATTTATTCAGATGCTCGATTTGGAAGAGACTT 2508  
Qy  
756 eLysLeuPheLysLysIlePheProSerLeuGluLeuAsnIleThrAspLeuLeuGluAs 776  
Db  
2509 TAAACTATTTAAAAAATTTTCTTCTCTGGAATTTAAATAAACAAGATTTACTTGAAGA 2568  
Qy  
776 pThrProArgGlnCysArgIleCysGlyGlyLeuAlaMetTyrGluCysArgGluCysTy 796  
Db  
2569 CACTCCCAGACAGTCCCGATATGAGAGGCTTCAATGATGATGATGATGATGATGATGATGAT 2628  
Qy  
796 rAspAspProAspIleSerAlaGlyLysIleLysGlnPheCysLysThrCysAsnThrG1 816  
Db  
2629 CGACCATTCGGACATCTCAGCTGGAAATAATCAAGCAGTTTGTAAAACTTCGCAACTCA 2688  
Qy  
816 nValHisLeuHisProLysArgLeuAsnHisLysTyrAsnProValSerLeuProLysAs 836  
Db  
2689 AGTCCACCTTCATCCGAAGAGGCTGAATCATATAATATACCAAGTGTCTACTTCCCCAAGA 2748  
Qy  
836 pleuProAspTrpAspTrpArgHisGlyCysIleProCysGlnAsnMetGluLeuPheAl 856

Db 2749 CTTACCCGACTGGGACTGGAGACAGCGCTGCATCCCTTCCGAGAAATATGAGTTATTGTC 2808  
Qy 856 aValLeuCyStleGluThrSerHisTyrrValAlaPheValLysTyrrGlyLysAspSe 876  
Db 2809 TGTTCCTTCGATAGAAACAGCCACTATGCTTTGTGAAATGATGGGAAGACCATTC 2868  
Qy 876 rAlaTrpLeuPhePheAspSerMetAlaAspArgAspGlyGlyGlnAsnGlyPheAsn1 896  
Db 2869 TGCCTGGCTCTCTTTTGACAGATGGCCGATCGGATGCTGCTCAGAAATGGCTTCAACAT 2928  
Qy 896 eProGlnValThrProCysProGluValGlyGluTyrrLeuLysMetSerLeuGluAspLe 916  
Db 2929 TCCTCAAGTCAACCCATGCCAGAGTAGGAGACTTGAAGATGCTCTGCGAAGACCT 2988  
Qy 916 uHisSerLeuAspSerArgArgIleGlnGlyCysAlaArgArgLeuLeuCysAspAlaTy 936  
Db 2989 GCATTCTTGGACTCCAGAGAAATCCAGGCTGTGCAGAAAGACTGCTTTGTGATGCATA 3048  
Qy 936 rMetCysMetTyrrGlnSerProThrMetSerLeuTyrrLys 949  
Db 3049 TATGTGCATGTACAGAGTCCAAACATGAGTTTGTACAAA 3088

RESULT 14

PCT-US02-27777-49

; Sequence 49: Application PC/TUS0227777

; GENERAL INFORMATION:

; APPLICANT: diadex, Inc.

; APPLICANT: Sun, Yongming

; APPLICANT: Liu, Chenghua

; APPLICANT: Salceda, Susana

; TITLE OF INVENTION: Compositions and Methods Relating to Breast Specific Genes and Pr

; FILE REFERENCE: DEX-0346

; CURRENT APPLICATION NUMBER: PCT/US02/27777

; CURRENT FILING DATE: 2002-08-29

; PRIOR APPLICATION NUMBER: US 60/316,306

; PRIOR FILING DATE: 2001-08-31

; NUMBER OF SEQ ID NOS: 170

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 49

; LENGTH: 6831

; TYPE: DNA

; ORGANISM: Homo sapien

PCT-US02-27777-49

Alignment Scores:

Pred. No.: 0 Length: 6831  
Score: 754.00 Matches: 948  
Percent Similarity: 98.54% Conservatives: 0  
Best Local Similarity: 98.54% Mismatches: 1  
Query Match: 79.45% Indels: 14  
DB: 1 Gaps: 0

US-09-671-687A-3 (1-949) x PCT-US02-27777-49 (1-6831)

Qy 1 MetSerSerGlyLeuTrpSerGlnGluLysValThrSerProTyrrTrpGluArgIle 20  
Db 1585 ATGAGTTTCAGGCTTATGGAGCCCAAGAAAAGTCACTTACCCTACTCTGGAGAGCGGANT 1644  
Qy 21 PheTyrrLeuLeuGlnGluCysSerValThrAspLysGlnThrClnLysLeuLeuLys 40  
Db 1645 TTTTACTTGTCTCTTCAAGATGCGCGTTACAGAAACAAACACAAAAGTCTCTTAAA 1704  
Qy 41 ValProLysGlySerIleGlnTyrrIleGlnAspArgSerValGlyHisSerArgIle 60  
Db 1705 GTACCGAAGGGAAGTATAGGACAGATATTTCAAGATCGTTCTGTGGGCAATTCAGGATT 1764  
Qy 61 ProSerAlaLysGlyLysLysAsnGlnIleGlyLeuLysIleLeuGlnProHisAla 80  
Db 1765 CCTTCTGCAAAAGGCAAGAAAATCAGATTGGATTAAAAATTCTAGAGCAACCTCATGCA 1824  
Qy 81 ValLeuPheValAspGlu---AspValValGluIleAsnGluLysPheThrGluLeuLeu 99

Db 1825 GTTCTCTTTGTTGATGAAAGGAGTGTGTAGAGATAAATGAAAGTTTCAAGAGTTACTT 1884  
Qy 100 LeuAlaIleThrAsnCysGluGluArgPheSerLeuPheLysAsnArgAsnArgLeuSer 119  
Db 1885 TTGGCAATTACCAATTTGTGAGGAGAGTTTCAGCTGTTTAAACACAGAAACAGACTAAGT 1944  
Qy 120 LysGlyLeuGlnIleAspValGlyCysProValLysValGlnLeuArgSerGlyGluGlu 139  
Db 1945 AAGGCTCCCAATAGACGTGGCTGCTCTGTGAAAGTACAGCTGAGATCTGGGGAAGAA 2004  
Qy 140 LysPheProGlyValValArgPheArgGlyProLeuLeuAlaGluArgThrValSerGly 159  
Db 2005 AAATTTCTTGGAGTTTACGCTTTCAGAGGACCCCTGTGTAGCAGAGAGACAGTCTCCGGA 2064  
Qy 160 IlePhePheGlyValGluLeuLeuGluGluArgGlyGlnGlyPheThrAspGlyVal 179  
Db 2065 ATATTCTTGGAGTTTGAATTTCTGGAAGAGAGTCTGCTGCTCAAGGTTTCACTGACGGGGTG 2124  
Qy 180 TyrrGlnGlyLysGlnLeuPheGlnCysAspGluAspCysGly---PheValAlaLeuAsp 198  
Db 2125 TACCAAGGGAACAGCTTTTTCAGTGTGATGAAGATTGTGGCTGTTTGTTCATTTGGAC 2184  
Qy 199 LysLeuGluLeuIleGluAspAspThrAlaLeuGluSerAspTyrrAlaGlyProGly 218  
Db 2185 AAGCTAGAACTCATAGAGATGATGACACTGCTATTTGAAAGTCAATTCACGAGGTCCTGGG 2244  
Qy 219 AspThrMetGlnValGluLeuProProLeuGluIleAsnSerArgValSerLeuLysGly 238  
Db 2245 GACACAAATCAGGCTGAACTTCTCTCTTGGAAATAAATCCAGAGTTTCTTTGAA--GGT 2303  
Qy 239 -GlyGluThrIleGluSerGlyThrValIlePheCysAspValLeuProGlyLysGluSe 258  
Db 2304 TGGAGAAAACATAGAAATCTGGAACAGTTATATCTGTGATGTTTGGCCAGAAAAGAAAG 2363  
Qy 258 rLeuGlyTyrrPheValGlyValAspMetAspAsnProIleGlyAsnTrpAspGlyArgph 278  
Db 2364 CTTAGGATATTTTGTGGTGTGACATGGATAACCTATTGGCAACTGGGATGGAAGATT 2423  
Qy 278 eAspGlyVal---LeuCysSerPheAlaCysValGluSerThrIleLeuLeuHisIleAs 297  
Db 2424 TGATGGAGTGCAGCTTTGTAGTTTTCGCTGTGTGAAAGTACAAATCTATTTCACATCAA 2483  
Qy 297 nAspIleIleProGluSerValThrGlnGluArgArgProProLysLeuAlaPheMetSe 317  
Db 2484 TGATATCATCCAGAGAGTGTGACGAGAAAAGAGGCGCTCCCAAACTTGCCCTTTATGTC 2543  
Qy 317 rArgGlyValGlyAspLysGlySerSerHisAsnLysProLysAlaThrGlySerTh 337  
Db 2544 AAGAGTGTGGGGACAAAGTTTCATCCAGTCAATATTAACCAAGGCTACAGGATCTAC 2603  
Qy 337 rSerAspProGlyAsnArg---ArgSerGluLeuPheTyrrThrLeuAsnGlySerSerVa 356  
Db 2604 CTCAGACCTCGGAATAGAAAACAGATCTGAATTTATTTTATACCTTAAATGGGCTCTCTGT 2663  
Qy 356 lAspSerGlnProGlnSerLysSerLysAsnThrTrpTyrrIleAspGluValAlaGluAs 376  
Db 2664 TGACTCACACCAACCAATCCAAATCAAAAATACATGTTACATTTGATGAAAGTTCAGAGA 2723  
Qy 376 pProAlaLysSerLeuThrGluIleSerThrAspPheAspArgSerSerProProLeuGlu 396  
Db 2724 CCTTGCAAAATCTCTTTACAGAGATATCTACAGACTTTTGACCGTTCTTCCACCACCTCCA 2783  
Qy 396 nProProProValAsnSerLeuThrThrGluAsnArgPheHisSerLeuProPheSerLe 416  
Db 2784 GCCTCTCTCTGTGAACCTCACTGACCAACGAGAACAGATCCCACTCTTTTACCATTCAGTCT 2843  
Qy 416 uThrLysMetProAsnThrAsnGlySerIleGlyHisSerProLeuSerLeuSerAlaGl 436  
Db 2844 CACCAAGATGCCCCAATACCAATGGAAGTATTGGCCACAGCTCCACTTTCTCTGTGAGCCCA 2903  
Qy 436 nSerValMetGluGluLeuAsnThrAlaProValGlnGluSerProProLeuAlaMetPr 456  
Db 2904 GTCTGTATGGAAGAGCTAAACACTGCAACCGCTCCAGAGAGAGTCCACCTTTGGCCATGCC 2963

QY 456 oProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLysGluAsnProPr 476  
Db 2964 TCCTGGGAACATCACATGCTCTAGAAAGTGGGCTCAATGGCTGAAGTTAAGGAGAACCCCTCC 3023  
QY 476 oPheTyGlyValIleArgTrrPileGlyGlnProProGlyLeuAsnGluValLeuAlaGI 496  
Db 3024 TTTCTATGGGGTAATCCCGTTGGATCGGTGAGCCAGCAGTGAATGAATGCTCGCTGG 3083  
QY 496 YLeuGluLeuGluAspGluCysAlaGlyCysThrAspGlyThrPheArgGlyThrArgTy 516  
Db 3084 ACTGGAACCTGGAGATGAGTGTGAGGCTGTACGATGGAACCTTCAGAGGCACCTCGGTA 3143  
QY 516 rPheThrCysAlaLeuLysLysAlaLeuPheValLysLeuLysSerCysArgProAspSe 536  
Db 3144 TTTCACTGTGCTTGAAGAGCGCTGTGTTGTGAACCTGAAGAGCTGCAGGCGCTGACTC 3203  
QY 536 rArgPheAlaSerLeuGlnProValSerAsnGlnIleGluArgCysAsnSerLeuAlaPh 556  
Db 3204 TAGGTTTGCATCATTTGCAGCGCGTTTCCAATCAGATTGAGCGCTGTAACCTCTTTAGCAT 3263  
QY 556 eGlyGlyTyrrLeuSerGluValValGluGluAsnThrProProLysMetGluLysGluGI 576  
Db 3264 TGGAGGCTACTTTAAGTGAAGTAGTAGAAGAAATACTCCACCAAAATGGAAAAAGAGG 3323  
QY 576 YLeuGluIleMetIleGlyLysLysLysGlyIleGlnGlyHisTyrrAsnSerCysTyrrLe 596  
Db 3324 CTTGGAGATATGATTGGGAGAGAGAGGCATCCAGGCTCATTAACATCTTTGTACTT 3383  
QY 596 uAspSerThrLeuPheCysLeuPheAlaPheSerSerValLeuAspThrValLeuLeuAr 616  
Db 3384 AGACTCAACCTTATTTCTGCTTATTTGCTTTAGTTCTGTCTGGACACTGTGTACTTAG 3443  
QY 616 gProLysGluLysAsnAspValGluTyrrTyrrSerGluThrGlnGluLeuLeuArgThrGI 636  
Db 3444 ACCCAAGAAAGACGATGTAGATATATTATAGTGAACCCCAAGAGCTACTGAGGACAGA 3503  
QY 636 uileValAsnProLeuArgIleTyrrGlyTyrrValCysAlaThrLysIleMetLysLeuAr 656  
Db 3504 AATTGTTAATCTCTGAGAATATATGATATGTGTGTCACCAAAATTTATGAACCTGAG 3563  
QY 656 gLysIleLeuGluLysValGluAlaAsrGlyPheThrSerGluGluLysAspProGI 676  
Db 3564 GAAATACTTGAAAGGTGGAGCTGTCATCAGATTTTACCTCTGAAGAAAAGATCCTGA 3623  
QY 676 uGluPheLeuAsnIleLeuPheHisIleLeuArgValGluProLeuLeuLysIleAr 696  
Db 3624 GGAATTTCTGAATATTTCTGTTTCATCATATTTTAAGGGTAGAACCTTTGCTAAAAATAAG 3683  
QY 696 gSerAlaGlyGlnLysValGlnAspCysTyrrPheTyrrGlnIlePheMetGluLysAsnGI 716  
Db 3684 ATCAGAGCTCAAAAGGTACAGATTTGTACTTCTATCAAAATTTTATGGAAAAAATGA 3743  
QY 716 uLysValGlyValProThrIleGlnGlnLeuLeuGluTrpSerPheIleAsnSerAsnLe 736  
Db 3744 GAAAGTTGGCGTTCACCAATTCACAGTGTAGAAATGGTCTTTTATCAACAGTAACCT 3803  
QY 736 uLysPheAlaGluAlaProSerCysLeuIleIleGlnMetProArgPheGlyLysAspPh 756  
Db 3804 GAAATTTGCAGAGGCACCATCATGCTGATTTATTCAGATGCTCGATTTTGGAAAAGACTT 3863  
QY 756 eLysLeuPheLysIlePheProSerLeuGluLeuAsnIleThrAspLeuLeuGluAs 776  
Db 3864 TAAACTATTTAAAAAATTTTTCTTCTCTGGAATTAATATAACAGATTTTACTTTGAAGA 3923  
QY 776 pThrProArgGlnCysArgIleCysGlyLeuAlaMetTyrrGluCysArgGluCysTy 796  
Db 3924 CACTCCCAGACAGTCCCGGATATGTGGAGGCTTGCATATGTATGATGTAGAGATGTCTA 3983  
QY 796 rAspAspProAspIleSerAlaGlyLysIleLysGlnPheCysLysThrCysAsnThrGI 816  
Db 3984 CGACGATCCGACATCTCAGCTGGAAATCAAGCAGTTTGTAAAACTTGCAACACTCA 4043

QY 816 nValHisLeuHisProLysArgLeuAsnHisLysTyrrAsnProValSerLeuProLysAs 836  
Db 4044 AGTCCACCTTCATCCGAAGAGGCTGAATCATATAATATAACCCAGTGTCACTTCCCAAGA 4103  
QY 836 pleuProAspTrpAspTrpArgHisGlyCysIleProCysGlnAsnMetGluLeuPheAl 856  
Db 4104 CTTACCCAGCTGGGACTGGAGACACGGCTGCATCCCTTCCAGAAATATGGAGTTATTTGC 4163  
QY 856 aValLeuCysIleGluThrSerHisTyrrValAlaPheValLysTyrrGlyLysAspAspSe 876  
Db 4164 TGTCTCTGCATAGAAACACAGCCACTATGTGTCTTTGTGAAGTATGGGAGACGATTTC 4223  
QY 876 rAlaTrpLeuPhePheAspSerMetAlaAspArgAspGlyGlyGlnAsnGlyPheAsnI 896  
Db 4224 TGCCTGGCTCTTCTTTGACAGCATGCCGATCGGATGGTGTGATGATGGCTTCAACAT 4283  
QY 896 eProGlnValThrProCysProGluValGlyGluTyrrLeuLysMetSerLeuGluAspLe 916  
Db 4284 TCCTCAAGTCAACCCATGCCAGAAAGTAGGAGAGTACTTTGAAGATGTCTCTGGAAGACCT 4343  
QY 916 uHisSerLeuAspSerArgArgIleGlnGlyCysAlaArgArgLeuLeuCysAspAlaTy 936  
Db 4344 GCATTTCTTGGACTCCAGAGAAATCCAAGGCTGTGCAGAAAGCTGCTTTGTGATGCATA 4403  
QY 936 rMetCysMetTyrrGlnSerProThrMetSerLeuTyrrLys 949  
Db 4404 TATGTGCATGTACCAGAGTCCACCAATGAGTTTGTACAAA 4443

## RESULT 15

PCT-US02-27777A-49  
; Sequence 49, Application PC/TUS0227777A  
; GENERAL INFORMATION:  
; APPLICANT: diadexus, Inc.  
; APPLICANT: Sun, Yongming  
; APPLICANT: Liu, Chenghua  
; APPLICANT: Salceda, Susana  
; TITLE OF INVENTION: Compositions and Methods related to Breast Specific Genes and Pro  
; FILE REFERENCE: DEX-0346  
; CURRENT APPLICATION NUMBER: PCT/US02/27777A  
; PRIOR FILING DATE: 2002-10-24  
; PRIOR APPLICATION NUMBER: US 60/316,307  
; PRIOR FILING DATE: 2001-08-31  
; NUMBER OF SEQ ID NOS: 170  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 49  
; LENGTH: 6831  
; TYPE: DNA  
; ORGANISM: Homo sapien  
PCT-US02-27777A-49

Alignment Scores:  
Pred. No.: 0 Length: 6831  
Score: 754.00 Matches: 948  
Percent Similarity: 98.54% Conservative: 0  
Best Local Similarity: 98.54% Mismatches: 1  
Query Match: 79.45% Indels: 14  
DB: 1 Gaps: 0

US-09-671-687A-3 (1-949) x PCT-US02-27777A-49 (1-6831)

QY 1 MetSerSerGlyLeuTrpSerGlnGluLysValThrSerProTyrrTrpGluuArgIle 20  
Db 1585 ATGAGTTGAGGCTTATGGAGCCAGAAAAAGTCACTTCACTTCTGAGAGAGCGGATT 1644  
QY 21 PheTyrrLeuLeuGlnGluCysSerValThrAspLysGlnThrGlnLysLeuLys 40  
Db 1645 TTTTACTTGTCTTCAAGAAATGACGCGTTACAGACAAACAAACAAAGCTCTCTTAAA 1704  
QY 41 ValProLysGlySerIleGlyGlnTyrrIleGlnAspArgSerValGlyHisSerArgIle 60  
Db 1705 GTACCGAAGGAGATATAGGACAGTATATTCAAGATCGTTCTGTGGGCAATTCAGAGATT 1764  
QY 61 ProSerAlaLysGlyLysLysAsnGlnIleGlyLeuLysIleLeuGluGlnProHisAla 80







```
Qy 796 rAspAspProAspIleSerAlaGlyLysIleLysGlnPheCysLysThrCysAsnThrG1 816
Db 3984 CGACGATCCGGACATCTCAGCTGGAAAAATCAAGCAGTTTGTAAACCTGCAACACTCA 4043
Qy 816 nValHisLeuHisProLysArgLeuAsnHisLysTyrAsnProValSerLeuProLysAs 836
Db 4044 AGTCCACCTTCATCCGAGAGGCTGAATCATATAATATAACCCAGGTGCACCTCCCAAGA 4103
Qy 836 pLeuProAspTrpAspTrpArgHisGlyCysIleProCysGlnAsnMetGluLeuPheAl 856
Db 4104 CTTACCCGACTGGGACTGGAGACACGGCTGCATCCCTTGCAGATATGGAGTTATTGGC 4163
Qy 856 aValLeuCysIleGluThrSerHisTyrValAlaPheValLysTyrGlyLysAspSe 876
Db 4164 TGTCTCTGCATAGAAACAAGCCACTATGTTGCTTTGTGAAGTATGGAGGACGATTC 4223
Qy 876 rAlaTrpLeuPhePheAspSerMetAlaAspArgAspGlyGlnAsnGlyPheAsnI1 896
Db 4224 TGCCTGGCTCTCTTTTGACAGCATGGCCGATCGGGATGGTGCAGATGGCTTCAACAT 4283
Qy 896 eProGlnValThrProCysProGluValGlyGluTyrLeuLysMetSerLeuGluAspLe 916
Db 4284 TCCTCAAGTCAACCCATGCCAGAAAGTAGGAGAGTCTTGAGATGTCTCTGGAGACCT 4343
Qy 916 uHisSerLeuAspSerArgArgIleGlnGlyCysAlaArgArgLeuLeuCysAspAlaTy 936
Db 4344 GCATTCTCTGGACTCCAGGAGAAATCCAAGGCTGTGCACGAGACTGCTTTGTGATGCATA 4403
Qy 936 rMetCysMetTyrGlnSerProThrMetSerLeuTyrLys 949
Db 4404 TATGTGATGTACCAGAGTCCAACAATGAGTTTGTACAAA 4443
```

Search completed: April 17, 2005, 13:44:38  
Job time : 7279 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: April 17, 2005, 07:25:28 ; Search time 590 Seconds  
(without alignments)

2489.941 Million cell updates/sec

Title: US-09-671-687A-3

Perfect score: 949

Sequence: 1 MSSGLWSQKVTSPYWEERI.....RLLCDAYMCYQSTMSLYK 949

Scoring table:

OLIGO  
Xgapop 60.0 , Xgapext 60.0  
Ygapop 60.0 , Ygapext 60.0  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 9152456 seqs, 774006986 residues

Word size: 1

Total number of hits satisfying chosen parameters: 18303315

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame+\_p2n.model -DEV=xlp  
-Q=/cgn2\_1/USPRO.spool/p/US09671687/runat\_15042005\_163040\_12041/app\_query.fasta\_1.1095  
-DB=Pending Patents NA New -QFMT=fastap -SUFFIX=Oligo.rnpn -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -SPART=1 -END=1 -MATRIX=oligo  
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=quality -THR MIN=1  
-ALIGN=15 -MODE=LOCAL -OUTFMT=pcio -NORW=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=2000000000 -USER=US09671687 @CGN 1.1.268 @runat\_15042005\_163040\_12041  
-NCFU=6 -ICFU=3 -NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100  
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : Pending Patents NA New.\*

1: /cgn2\_6/ptodata/2/pna/PCT NEW COMB.seq.\*  
2: /cgn2\_6/ptodata/2/pna/PCT NEW COMB.seq2.\*  
3: /cgn2\_6/ptodata/2/pna/US06 NEW COMB.seq.\*  
4: /cgn2\_6/ptodata/2/pna/US07 NEW COMB.seq.\*  
5: /cgn2\_6/ptodata/2/pna/US08 NEW COMB.seq.\*  
6: /cgn2\_6/ptodata/2/pna/US09 NEW COMB.seq.\*  
7: /cgn2\_6/ptodata/2/pna/US10 NEW COMB.seq.\*  
8: /cgn2\_6/ptodata/2/pna/US11 NEW COMB.seq.\*  
9: /cgn2\_6/ptodata/2/pna/US11 NEW COMB.seq2.\*  
10: /cgn2\_6/ptodata/2/pna/US11 NEW COMB.seq2.\*  
11: /cgn2\_6/ptodata/2/pna/US11 NEW COMB.seq3.\*  
12: /cgn2\_6/ptodata/2/pna/US11 NEW COMB.seq3.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	685	72.2	5371	7	US-10-760-678-3
2	641	67.5	2523	7	US-10-122-851-290
C 3	86	9.1	1160	8	US-10-450-763-26328
C 4	27	2.8	600	7	US-10-972-079-20538
C 5	27	2.8	600	7	US-10-972-079-20539
C 6	27	2.8	600	7	US-10-972-079-20540

Sequence 19097, A	7	US-10-972-079-19097	600	0.9	9
Sequence 114267, A	8	US-60-655-875-114267	1003	0.9	9
Sequence 75550, A	C 9	US-60-655-875-75550	1854	0.9	9
Sequence 17891, A	9	US-60-655-875-17891	2900	0.9	9
Sequence 58260, A	C 11	US-60-655-875-58260	3949	0.9	9
Sequence 1731, A	C 12	US-60-655-875-1731	6161	0.9	9
Sequence 5500, App	C 13	US-60-655-875-5500	6758	0.9	9
Sequence 956, App	14	US-10-287-436A-956	149034	0.9	14
Sequence 9068, A	8	US-11-036-317-90968	25	0.8	25
Sequence 105317, A	8	US-11-036-317-105317	25	0.8	25
Sequence 118928, A	8	US-11-036-317-118928	25	0.8	25
Sequence 130543, A	8	US-11-036-317-130543	25	0.8	25
Sequence 1033757, A	C 19	US-10-605-924-1033757	28	0.8	28
Sequence 750, App	20	PCT-IB03-06509-750	201	0.8	201
Sequence 15680, A	8	US-60-659-397-15680	201	0.8	201
Sequence 66300, A	8	US-60-655-875-66300	274	0.8	274
Sequence 97836, A	8	US-60-655-875-97836	325	0.8	325
Sequence 92236, A	8	US-10-972-079-92236	599	0.8	599
Sequence 45369, A	C 25	US-10-972-079-45369	600	0.8	600
Sequence 45370, A	C 26	US-10-972-079-45370	600	0.8	600
Sequence 45371, A	C 27	US-10-972-079-45371	600	0.8	600
Sequence 91224, A	28	US-10-972-079-91224	600	0.8	600
Sequence 2961, App	C 29	US-11-033-545-2961	601	0.8	601
Sequence 4580, App	30	US-11-033-545-4580	601	0.8	601
Sequence 9356, App	C 31	US-11-033-545-9356	601	0.8	601
Sequence 533, App	C 32	US-10-932-182A-533	669	0.8	669
Sequence 70912, A	33	US-60-655-875-70912	695	0.8	695
Sequence 3210, App	34	US-11-031-175-3210	771	0.8	771
Sequence 37846, A	C 35	US-60-655-875-37846	952	0.8	952
Sequence 109221, A	C 36	US-60-655-875-109221	982	0.8	982
Sequence 108648, A	37	US-60-655-875-108648	1003	0.8	1003
Sequence 108901, A	38	US-60-655-875-108901	1003	0.8	1003
Sequence 52287, A	39	US-60-655-875-52287	1089	0.8	1089
Sequence 1798, App	C 40	US-11-079-463-1798	1092	0.8	1092
Sequence 1798, App	C 41	US-11-079-045-1798	1092	0.8	1092
Sequence 77350, A	C 42	US-60-655-875-77350	1162	0.8	1162
Sequence 37150, A	C 43	US-60-655-875-37150	1183	0.8	1183
Sequence 6149, App	C 44	US-60-655-875-6149	1237	0.8	1237
Sequence 48561, A	C 45	US-60-655-875-48561	1286	0.8	1286

#### ALIGNMENTS

##### RESULT 1

US-10-760-678-3  
; Sequence 3, Application US/10760678  
; GENERAL INFORMATION:  
; APPLICANT: Derry, Jonathan  
; APPLICANT: Farnlow, William  
; APPLICANT: Dougall, William  
; TITLE OF INVENTION: SCREENING ASSAYS FOR AGONISTS OR ANTAGONISTS OF CD40 SIGNALING  
; FILE REFERENCE: 3198  
; CURRENT APPLICATION NUMBER: US/10/760,678  
; CURRENT FILING DATE: 2004-01-20  
; PRIOR APPLICATION NUMBER: US/09/851,673  
; PRIOR FILING DATE: 2001-05-08  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: Patentin version 3.0  
; SEQ ID NO 3  
; LENGTH: 5371  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (392)..(3262)  
US-10-760-678-3

Alignment Scores:  
Pred. No.: 0  
Score: 685.00  
Percent Similarity: 97.63%  
Best Local Similarity: 97.63%  
Query Match: 72.18%  
Length: 5371  
Matches: 948  
Conservative: 0  
Mismatches: 1  
Indels: 23

DB:	7	Gaps:	0
US-09-671-687A-3 (1-949) x US-10-760-678-3 (1-5371)			
QY, 1 MetSerSerGlyLeuTyrSerGlnGluLysValThrSerProTyrTrpGluLysArgile 20			
DB 392 ATGAGTTTCAGGCTTATGGAGGCAAGAAAGTCACTTCACCCCTACTGGGAAGAGCGGATT 451			
QY 21 PheTyrLeuLeuLeuGlnGluCysSerValThrAspLysGlnThrGlnLysLeuLeuLys 40			
DB 452 TTTTACTTGTCTTCAGANTGACGCTTACAGACAACAACAACAAGAGCTCTTAAA 511			
QY 41 ValProLysGlySerIleGlyGlnTyrIleGlnAspArgSerValGlyHisSerArgile 60			
DB 512 GTACCGAAGGAAGTATAGGACAGTATATTCAAGATCGTTCTGTGGGCACTTCAAGGATT 571			
QY 61 ProSerAlaLysGlyLysLysAsnGlnIleGlyLeuLysIleLeuGluGlnProHisAla 80			
DB 572 CCTTCTGCAAAAGGCAAGAAAATCAGATTGGATTAAAAAATTTCTAGAGCAACCTCATGCA 631			
QY 81 ValLeuPheValAspGlu---AspValValGluIleAsnGluLysPheThrGluLeuLeu 99			
DB 632 GTTCTCTTGTGATGAAGAAGATGTTGTAGAGATAAATGAATAAGTTTACAGAGTTACTT 691			
QY 100 LeuAlaIleThrAsnCysGluGluArgPheSerLeuPheLysAsnArgAsnArgLeuSer 119			
DB 692 TTGGCAATTACCAATTGTGAGGAGAGGTTTCAGCCTGTTTAAAAACAGAAAACAGACTAAGT 751			
QY 120 LysGlyLeuGlnIleAspValGlyCysProValLysValGlnLeuArgSerGlyGluGlu 139			
DB 752 AAAGCGCTTCCAAATAGACGTGGGTCTCTGTGAAAGTACAGCTGAGATCTGGGGAGAA 811			
QY 140 LysPheProGlyValValArgPheArgGlyProLeuLeuAlaGluArgThrValSerGly 159			
DB 812 AATTTCTTGAGTTGATCGCTTACAGAGACCCCTGTTAGCAGAGAGACAGTCTCCGGA 871			
QY 160 IlePhePheGlyValGluLeuGluGluArgGlyArgGlyGlnGlyPheThrAspGlyVal 179			
DB 872 ATATTTCTTGAGTTGAATGCTGGAAGAAGTCTGTGTCAAGTTTCACTGACGGGTG 931			
QY 180 TyrGlnGlyLysGlnLeuPheGlnCysAspGluAspCysGly---PheValAlaLeuAsp 198			
DB 932 TACCAAGGGAACAGCTTTTTCAGTGTGATGAAGATTCTGGCGTGTGTTGTTGTCATTGGAC 991			
QY 199 LysLeuGluLeuIleGluAspAspThrAlaLeuGluSerAspTyrAlaGlyProGly 218			
DB 992 AAGCTAGAACTCATAGAAAGATGATGACACTGCTTGGAAAGTGATTACGCAAGTCTCTGG 1051			
QY 219 AspThrMetGlnValGluLeuProProLeuGluIleAsnSerArgValSerLeuLysGly 238			
DB 1052 GACACATGCAAGTTCGAATCTCTCTTGGAAATAACTCCAGAGTTTCTTTGAA--GGT 1110			
QY 239 -GlyGluThrIleGluSerGlyThrValIlePheCysAspValLeuProGlyLysGluSe 258			
DB 1111 TGGAGAAACAATAAGAACTGGAACAGTTATTTCTGTGATGTTTTCAGGAAAGAAAG 1170			
QY 258 rLeuGlyTyrPheValGlyValAspMetAspAsnProIleGlyAsnTrpAspGlyArgPh 278			
DB 1171 CTTTAGGATATTTTGTGTGTGGACATGATTAACCTTATGGCACTGGGATGGAAGATT 1230			
QY 278 eAspGlyVal---LeuCysSerPheAlaCysValGluSerThrIleLeuLeuHisAla 297			
DB 1231 TGATGGAGTGAGCTTTGTAGTTTTCGTGTGTTGAAAGTACAAATCTATTGCAATCAA 1290			
QY 297 nAspIleIlePro-----GluSerValThrGlnGluArgArgProProLysLeuAl 314			
DB 1291 TGATATCATCCAGCTTTTATCAGAGAGTGTGACGAGGAGAGGAGGCTCCCAACTTGC 1350			
QY 314 aPheMetSerArgGlyValGlyAspLysGlySerSerHisAsnLysProLysAlaTh 334			
DB 1351 CTTTATGTCAAGAGGTGTGGGACAAAGGTTTCATCCAGTCAATAAACAAGGCTAC 1410			
QY 334 rGlySerThrSerAspProGlyAsnArg---ArgSerGluLeuPheTyrThrLeuAsnG 353			
DB 1411 AGGACTACCTCAGACCCCTGGAAATAGAAAACAGATCTGAATATTATTTATACCTTAAATCG 1470			
QY 353 ySerSerValAspSerGlnProGlnSerLysSerLysAsnThrTyrTyrIleAspGluVa 373			
DB 1471 GTCTTCTGTGACTCACAAACCAATCAAAATCAAAAATCATGGTACATTGATGAAGT 1530			
QY 373 lAlaGluAspProAlaLysSerLeuThrGluIleSerThrAspPheAspArgSerSerPr 393			
DB 1531 TGCAGAAAGCCCTGCAAAATCTCTACAGAGATATCTACAGACTTTTGACCGTCTCTCAC 1590			
QY 393 oProLeuGlnProProValAsnSerLeuThrThrGluAsnArgPheHisSerLeuPr 413			
DB 1591 ACCACTCAGCCTCTCTGTGAACCTCACTGACCAACGAGAACAGATCCCACTCTTTACC 1650			
QY 413 oPheSerLeuThrLysMetProAsnThrAsnGlySerIleGlyHisSerProLeuSerLe 433			
DB 1651 ATTCAGTCTCACCAGATGCCCAATACCAATGGAAAGTATTGGCCACAGTCCACTTTCTCT 1710			
QY 433 uSerAlaGlnSerValMetGluLeuLeuAsnThrAlaProValGlnGluSerProProLe 453			
DB 1711 GTCAGCCAGTCTGTAAATGGAAAGACTAAACACTGCAACCCGTCGAAGAGAGTCCACCTT 1770			
QY 453 uAlaMetProProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLysG 473			
DB 1771 GGCCATGCTCTCTGGGAACCTCACATGGTCTAGAAGTGGCTCATTTGGCTGAAGTTAAGCA 1830			
QY 473 uAsnProProPheTyrGlyValIleArgTrpIleGlyGlnProProGlyLeuAsnGluVa 493			
DB 1831 GAACCTCTCTTCTATGGGGTAAATCCGTTGGATCGTGCAGCCAGGACTGAATGAAGT 1890			
QY 493 lLeuAlaGlyLeuGluLeuGluAspGluCysAlaGlyCysThrAspGlyThrPheArgG 513			
DB 1891 GCTCTCGACTGGAACTGGAAAGATGATGTCAGGCTGTACGGATGGAACCTTCAGAGG 1950			
QY 513 yThrArgTyrPheThrCysAlaLeuLysLysAlaLeuPheValLysLeuLysSerCysAr 533			
DB 1951 CACTCGGTATTTCACTGTGCCCTGAAGAAGCGCTGTTGTGTGAATCTGAAGAGCTGCAG 2010			
QY 533 gProAspSerArgPheAlaSerLeuGlnProValSerAsnGlnIleGluArgCysAsnSe 553			
DB 2011 GCCTGACTCTAGGTTTGCATCATTCAGCCGGTTTCCAATCAGATTGAGCGCTGTAACTC 2070			
QY 553 rLeuAlaPheGlyGlyTyrLeuSerGluValValGluGluAsnThrProProLysMetG 573			
DB 2071 TTTGACATTTGGAGGCTACTTAAAGTGAAGTAGTAGAAGAAATACTCCACCAAAATGGA 2130			
QY 573 uLysGluGlyLeuGluIleMetIleGlyLysLysLysGlyIleGlnGlyHisTyrAsnSe 593			
DB 2131 AAAAGAAGGCTTGGAGATATTAATTTGGGAAGAAAGAAAGCATCCAGGGTCATTACNAATC 2190			
QY 593 rCysTyrLeuAspSerThrLeuPheCysLeuPheAlaPheSerSerValLeuAspThrVa 613			
DB 2191 TTGTTACTTAGACTCAACCTTATTCTGCTTATTGCTTTTAGTCTTGTCTGACACTGT 2250			
QY 613 lLeuLeuArgProLysGlyLysAsnAspValGluTyrTyrSerGluThrGlnGluLeuSe 633			
DB 2251 GTTACTTAGACCCAAAGAAAGAAAGCGATGTAGAATATTATAGTGAACCCCAAGAGCTACT 2310			
QY 633 uArgThrGluIleValAsnProLeuArgIleTyrGlyTyrValCysAlaThrLysIleMe 653			
DB 2311 GAGGACAAATTTGTAATCTCTCAGAAATATATGGAATATGTTGTGCCCAAAAATTTAT 2370			
QY 653 tLysLeuArgLysIleLeuGluLysValGluAlaAlaSerGlyPheThrSerGluGluLys 673			
DB 2371 GAAACTGAGGAAATTAATTTGAAAAGGTGGAGGCTGCATCAGGATTTACCTCTGAAGAAA 2430			
QY 673 aAspProGluGluPheLeuAsnIleLeuPheHisIleLeuArgValGluProLeuLe 693			
DB 2431 AGATCCTCAGGAAATTTCTTGAATATTCTGTTCATCATATTTTAAAGGTAGAACCTTTGCT 2490			
QY 693 uLysIleArgSerAlaGlyGlnLysValGlnAspCysTyrPheTyrGlnIlePheMetG 713			

```
Db 2491 AAAAATAGATCAGCAGGTCAAAAGGTACAGATTGTCTCTATCAAAATTTTATGGA 2550
Qy 713 uLysAsnGluLysValGlyValProThrIleGlnLeuLeuGluTTPSerPheIleAs 733
Db 2551 AAAAATGAGAAAGTTGGCGTTCACCAATTCAGCAGTTGTAGAAATGCTCTTTATCAA 2610
Qy 733 nSerAsnLeuLysPheAlaGluAlaProSerCysLeuIleIleGlnMetProArgPheGl 753
Db 2611 CAGTAACCTGAAATTTGCAGAGGCCACCATCATGCTCTGATATTTCAGATGCCCTCGATTTGG 2670
Qy 753 yLysAspPheLysLeuPheLysLysIlePheProSerLeuGluLeuAsnIleThrAspie 773
Db 2671 AAAAGACTTTAAACTATTATAAAAAATTTTCTCTCTCGGAATTAATAATATACAGATT 2730
Qy 773 uLeuGluAspThrProArgGlnCysArgIleCysGlyGlyLeuAlaMetTyrGluCysAr 793
Db 2731 ACTTGAGACACTCCAGACAGTGGCGGATATGTGAGGGCTTGCAATGATAGTGATG 2790
Qy 793 gGluCysTyrAspAspProAspIleSerAlaGlyLysIleLysGlnPheCysLysThrCy 813
Db 2791 AGAATGCTACGACGATCCGGACATCTCAGCTGGAATAATCAAGCAGTTTGTAAAACTG 2850
Qy 813 sAsnThrGlnValHisLeuHisProLysArgLeuAsnHisLysTyrAsnProValSerIe 833
Db 2851 CAACACTCAAGTCCACCTTCATCCGAGAGGCTGAATCATATAATATAACCCAGTGTCACT 2910
Qy 833 uProLysAspLeuProAspTTPAspTTPAspTTPArgHisGlyCysIleProCysGlnAsnMetGl 853
Db 2911 TCCCAAGAGACTTACCCGACTGGGACTGGAGACACGGCTGCATCCCTTGCAGAAATATGGA 2970
Qy 853 uLeuPheAlaValLeuCysIleGluThrSerHisTyrValAlaPheValLysTyrGlyIy 873
Db 2971 GTTATTTGCTGTCTCTGCATAGAAACAGCCACTATGTGTTTGTGAAGTATGGGAA 3030
Qy 873 sAspAspSerAlaTrpLeuPhePheAspSerMetAlaAspArgAspGlyGlyGlnAsnGl 893
Db 3031 GGACGATTTCTGCTGCTCTTTGACAGCATGGCCGATCGGATGGTGGTGCAGAAATGG 3090
Qy 893 yPheAsnIleProGlnValThrProCysProGluValGlyGluTyrLeuLysMetSerIe 913
Db 3091 CTTCAACATTTCTCAAGTCACCCCATGCCAGAGTAGGAGAGTACTTGAAGATGTCTCT 3150
Qy 913 uGluAspLeuHisSerLeuAspSerArgArgIleGlnGlyCysAlaArgArgLeuLeuCy 933
Db 3151 GGAAGACTGCTGATCTCTTGAGCTCCAGAGAAATCCAGGCTGTGCACGAAGACTGCTTTG 3210
Qy 933 sAspAlaTyrMetCysMetTyrGlnSerProThrMetSerLeuTyrLys 949
Db 3211 TGATGCATATATGTGATGTACCATGACAGTCCAACATGAGTTGTACAAA 3259
```

## RESULT 2

```
US-10-122-851-290
; Sequence 290, Application US/10122851
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Ren, Feiyan
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: Novel Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784C1P2BDV3
; CURRENT APPLICATION NUMBER: US/10/122,851
; CURRENT FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: 09/620,312
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1104
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 290
```

```
; LENGTH: 2523
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (81)..(2138)
; US-10-122-851-290
Alignment Scores:
Pred. No.: 0 Length: 2523
Score: 641.00 Matches: 668
Percent Similarity: 99.55% Conservative: 0
Best Local Similarity: 99.55% Mismatches: 0
Query Match: 67.54% Indels: 3
DB: 7 Gaps: 0
US-09-671-687A-3 (1-949) x US-10-122-851-290 (1-2523)
```

```
Qy 282 LeuCysSerPheAlaCysValGluSerThrIleLeuLeuHisIleAsnAspIlellePro 301
Db 129 CTTTGTAGTTTGGTGTGTTGAAAGTACAATTTCTATTGCACATCAATGATATCATCCCA 188
Qy 302 GluSerValThrGlnGluArgProProLysLeuAlaPheMetSerArgGlyValGly 321
Db 189 GAGAGTGTGACGAGAAAGAGGCTCCCAAACTTGCCTTTATGTCAAGAGGTGTGGG 248
Qy 322 AspLysGlySerSerHisAsnLysProLysAlaThrGlySerThrSerAspProGly 341
Db 249 GACAAGGTTTTCAGTCAGTCATATAAACCAAGAGGTACAGGATCTACCTCAGACCTTGG 308
Qy 342 AsnArg---ArgSerGluLeuPheTyrThrLeuAsnGlySerSerValAspSerGlnPro 360
Db 309 ATAGAAACAGATCTGAATTTATTTATACCTTAATGGGTCTTCTGTGACTCACAACCA 368
Qy 361 GlnSerLysSerLysAsnThrTrpTyrIleAspGluValAlaGluAspProAlaLysSer 380
Db 369 CAATCCAAATCAAAAAATACATGTTGATGAAAGTTGCAGAGACCCCTGCAAAATCT 428
Qy 381 LeuThrGluIleSerThrAspPheAspArgSerSerProProLeuGlnProProVal 400
Db 429 CTTACAGAGATATCTACAGACTTTGACCGTTCTTCACCAACCACTCCAGCCCTCTCTGTG 488
Qy 401 AsnSerLeuThrThrGluAsnArgPheHisSerLeuProPheSerLeuThrLysMetPro 420
Db 489 AACTCTACTGACCAACCGAGAACAGATTCACCTCTTTACCAATTCACTCTCACAAGATGCC 548
Qy 421 AsnThrAsnGlySerIleGlyHisSerProLeuSerLeuSerAlaGlnSerValMetGlu 440
Db 549 AATACCAATGGAAGTATTGGCCACAGTCACCTTCTCTGTGAGCCAGCTCTGTAATGAA 608
Qy 441 GluLeuAsnThrAlaProValGlnGluSerProProLeuAlaMetProProGlyAsnSer 460
Db 609 GAGCTAAACACTGCACCCCTCCAAGAGAGTCCACCCCTTGCCCATGCTCTCTGGAACCTCA 668
Qy 461 HisGlyLeuGluValGlySerLeuAlaGluValLysGluAsnProProPheTyrGlyVal 480
Db 669 CATGGTGTAGAAGTGGGCTCATTTGGCTGAAGTTAAGAGAACCCCTCTTCTATGGGGA 728
Qy 481 IleArgTrpIleGlyGlnProProGlyLeuAsnGluValLeuAlaGlyLeuGluLeuGlu 500
Db 729 ATCCGTTGGATCGGTGAGCAACAGGACTGAATGAAGTGTCTCGCTGAGCTGGAACTGAA 788
Qy 501 AspGluCysAlaGlyCysThrAspGlyThrPheArgGlyThrArgTyrPheThrCysAla 520
Db 789 GATGAGTGTGACGGCTGTACGGATGGAACCTTCAGAGGCACCTCGGTATTTTCACTGTGCC 848
Qy 521 LeuLysLysAlaLeuPheValLysLeuLysSerCysArgProAspSerArgPheAlaSer 540
Db 849 CTGAAGAAGCGCTGTGTTGTGAACACTGAAGAGCTGACGGCTGACTCTAGTTTGTGATCA 908
Qy 541 LeuGlnProValSerAsnGlnIleGluArgCysAsnSerLeuAlaPheGlyGlyTyrLeu 560
Db 909 TTGCAGCCGGTTTCCATCAGATTGAGCGCTGTAACCTCTTAGCATTTGGAGGCTACTTGA 968
```

```
Qy 561 SerGluValValGluGluAsnThrProProLysMetGluLysGluGluLeuMet 580
Db 969 AGTGAAGTAGTAGAAGAAATACTCCACCAAAATGGAAGAGGCTGGAGATAATG 1028
Qy 581 IleGlyLysLysGlyIleGlnGlyHisTyrAsnSerCysTyrLeuAspSerThrLeu 600
Db 1029 ATTGGGAAGAAGAAGGATCAGGGTCATTACAATCTTGTACTTAGACTCAACCTTA 1088
Qy 601 PheCysLeuPheAlaPheSerSerValLeuAspThrValLeuLeuArgProLysGluLys 620
Db 1089 TTCTGCTATTGCTTTTGTAGTTCTGTTCTGGACACTGTGTACTTAGACCCAAAGAAAG 1148
Qy 621 AsnAspValGluTyrTyrSerGluThrGlnGluLeuLeuArgThrGluLeuValAsnPro 640
Db 1149 AACGATGTAGAATATTATAGTGAACCCCAAGAGCTACTGAGCAGAGAATTTGTAATCT 1208
Qy 641 LeuArgIleTyrGlyTyrValCysAlaThrLysIleMetLysLeuArgLysIleLeuGlu 660
Db 1209 CTGGAATATATGGATATGTGTGTGCCCAAAATTTATGAACCTGAGGAAATACTTGAA 1268
Qy 661 LysValGluAlaAlaSerGlyPheThrSerGluGluLysAspProGluGluPheLeuAsn 680
Db 1269 AAGTGGAGGCTGCATCAGGATTTACCTCTGAAGAAAGATCCTGAGGAATCTTGAT 1328
Qy 681 IleLeuPheHisHisIleLeuArgValGluProLeuLeuLysIleArgSerAlaGlyGln 700
Db 1329 ATTCTGTTTCATCATATTATTAAGGCTAGAACCTTTTGCTAAAAATTAAGATCAGCAGTCA 1388
Qy 701 LysValGlnAspCysTyrPheTyrGlnIlePheMetGluLysAsnGluLysValGlyVal 720
Db 1389 AAGGTACAAGATTGTACTTCTATCAAAATTTTATGGAAGAAATGAGAAAGTTGGCGTT 1448
Qy 721 ProThrIleGlnGlnLeuLeuGluTrpSerPheIleAsnSerAsnLeuLysPheAlaGlu 740
Db 1449 CCCCAATTACGAGTGTAGATGGTCTTTTATCAACAGTAACTGAAATTTGCAGAG 1508
Qy 741 AlaProSerCysLeuIleGlnMetProArgPheGlyLysAspPheLysLeuPheLys 760
Db 1509 GCACCATCATGCTGATATTATTCAGATGCTCGATTTGGNAAGACTTTAACTATTAAA 1568
Qy 761 LysIlePheProSerLeuGluLeuAsnIleThrAspLeuLeuGluAspThrProArgGln 780
Db 1569 AAAATTTTCTCTCTCGAATTAATATAACAGATTTTACTTGAAGACACTCCCGACAG 1628
Qy 781 CysArgIleCysGlyLysAlaMetTyrGluCysArgGluCysTyrAspAspProAsp 800
Db 1629 TGCCGGATATGTGGAGGCTTGCAATGATGATGATGATGATGATGATGATGATGATGAT 1688
Qy 801 IleSerAlaGlyLysIleLysGlnPheCysLysThrCysAsnThrGlnValHisLeuHis 820
Db 1689 ATCTCAGCTGGAANAATCAAGAGTTTGTAAACCTGCAACACTCAAGTCCACCTTCAT 1748
Qy 821 ProLysArgLeuAsnHisLysTyrAsnProValSerLeuProLysLeuAspLeuProAspTrp 840
Db 1749 CCGAAGAGCTGATCATATAATATAACCCAGTGTCACTTCCCAAGACTTACCCGACTGG 1808
Qy 841 AspTrpArgHisGlyCysIleProCysGlnAsnMetGluLeuPheAlaValLeuCysIle 860
Db 1809 GACTGGAGACACGGCTGCTCCCTGCCAGAAATATGGAGTTATTGCTGCTCTGTCATA 1868
Qy 861 GluThrSerHisTyrValAlaPheValLysTyrGlyLysAspSerAlaTrpLeuPhe 880
Db 1869 GAAACAACCACTATGTTGCTTTTGTGAAGTATGGGAAGGAGATTCGCCCTGGCTCTTC 1928
Qy 881 PheAspSerMetAlaAspArgAspGlyGlyGlnAsnGlyPheAsnIleProGlnValThr 900
Db 1929 TTTGACACATGGCCGATCGGATGTTGTTGATGATGATGATGATGATGATGATGATGAT 1988
Qy 901 ProCysProGluValGlyGluTyrLeuLysMetSerLeuGluAspLeuHisSerLeuAsp 920
Db 1989 CCATGCCCAGAGTAGGAGAGTACTTGAAGATGCTCTCTGGAAGACCTGCAATTCCTTGGAC 2048
```

```
Qy 921 SerArgArgIleGlnGlyCysAlaArgArgLeuLeuCysAspAlaTyrMetCysMetTyr 940
Db 2049 TCCAGGAGATCCCAAGGCTGTGCACCAAGACTGCTTTGTGATGCATATATGTGCATGTAC 2108
Qy 941 GlnSerProThrMetSerLeuTyrLys 949
Db 2109 CAGAGTCCAAACAATGAGTTGTACAAA 2135
```

## RESULT 3

```
US-10-450-763-26328/c
; Sequence 26328, Application US/10450763
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 7901P3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 26328
; LENGTH: 1160
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIMILAR
; LOCATION: 1639..167
; OTHER INFORMATION: 96% homologous to Homo sapiens Familial Cylinidromatosis
; OTHER INFORMATION: Gene, accession number AJ250014, Smith-Waterman Score=963.
US-10-450-763-26328
```

## Alignment Scores:

```
Pred. No.: 8, 02e-76 Length: 1160
Score: 86.00 Matches: 86
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 9.06% Indels: 0
DB: 8 Gaps: 0
```

US-09-671-687A-3 (1-949) x US-10-450-763-26328 (1-1160)

```
Qy 1 MetSerSerGlyLeuTrpSerGlnGluLysValThrSerProTyrTrpGluGluArgile 20
Db 979 ATGAGTTTCAGGCTTATGGAGCCCAAGAAAAGTCACTTCACCTACTTGGGAGAGCGGATT 920
Qy 21 PheTyrLeuLeuLeuGlnGluCysSerValThrAspLysGlnThrGlnLysLeuLys 40
Db 919 TTTTACTTGTCTTCAAGAAATGACGCGTTACAGACAAACCAACAAAGCTCTTTAA 860
Qy 41 ValProLysGlySerIleGlyGlnTyrIleGlnAspArgSerValGlyHisSerArgile 60
Db 859 GTACCGAAGGGAAGTATAGGACAGTATATTCAAGATCGTCTTGTGGGCAATCAAGGATT 800
Qy 61 ProSerAlaLysGlyLysLysAsnGlnIleGlyLeuLysIleLeuGluGlnProHisAla 80
Db 799 CCTTCTGCAAAAGGCAAGAAAATCAGATTGGATTAAAAATTTAGAGCAACCTCATGCA 740
Qy 81 ValLeuPheValAspGlu 86
Db 739 GTTCTCTTTTGTGATGA 722
```

## RESULT 4

```
US-10-972-079-20538/c
; Sequence 20538, Application US/10972079
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: Denise Sue K.
; APPLICANT: ROSENFIELD, David
```

APPLICANT: KERR, Richard  
APPLICANT: BATES, Stephen  
APPLICANT: HOLM, Tom  
TITLE OF INVENTION: METHODS & SYSTEMS FOR INFERRING TRAITS TO BREED & MANAGE NON-BEER  
TITLE OF INVENTION: LIVESTOCK  
FILE REFERENCE: MM1110-2  
CURRENT APPLICATION NUMBER: US/10/972,079  
CURRENT FILING DATE: 2004-10-22  
PRIOR APPLICATION NUMBER: US 60/514,333  
PRIOR FILING DATE: 2003-10-24  
NUMBER OF SEQ ID NOS: 96631  
SOFTWARE: Patentin version 3.1  
SEQ ID NO 20538  
LENGTH: 600  
TYPE: DNA  
ORGANISM: Chicken 19866894231145\_1  
US-10-972-079-20538

Alignment Scores:  
Pred. No.: 1.15e-16 Length: 600  
Score: 27.00 Matches: 27  
Percent Similarity: 100.00% Conservatives: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 2.85% Indels: 0  
DB: 7 Gaps: 0

US-09-671-687A-3 (1-949) x US-10-972-079-20538 (1-600)

QY 714 LysAsnGluLysValGlyValProThrIleGlnLeuLeuGluTrpSerPheIleAsn 733  
|||||  
Db 396 AAAAAAGAAAGTTGGAGTCCGACAATYCACGAGTTACTGGAGTGGTCTTCATCAAC 337  
|||||

QY 734 SerAsnLeuLysPheAlaGlu 740  
Db 336 AGCAACTTGAAGTTTGCAGAG 316  
|||||

RESULT 5  
US-10-972-079-20539/c  
Sequence 20539, Application US/10972079  
GENERAL INFORMATION:  
APPLICANT: MMI GENOMICS, INC.  
APPLICANT: DENISE, Sue K.  
APPLICANT: ROSENFELD, David  
APPLICANT: KERR, Richard  
APPLICANT: BATES, Stephen  
APPLICANT: HOLM, Tom  
TITLE OF INVENTION: METHODS & SYSTEMS FOR INFERRING TRAITS TO BREED & MANAGE NON-BEER  
TITLE OF INVENTION: LIVESTOCK  
FILE REFERENCE: MM1110-2  
CURRENT APPLICATION NUMBER: US/10/972,079  
CURRENT FILING DATE: 2004-10-22  
PRIOR APPLICATION NUMBER: US 60/514,333  
PRIOR FILING DATE: 2003-10-24  
NUMBER OF SEQ ID NOS: 96631  
SOFTWARE: Patentin version 3.1  
SEQ ID NO 20539  
LENGTH: 600  
TYPE: DNA  
ORGANISM: Chicken 19866894231145\_2  
US-10-972-079-20539

Alignment Scores:  
Pred. No.: 1.15e-16 Length: 600  
Score: 27.00 Matches: 27  
Percent Similarity: 100.00% Conservatives: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 2.85% Indels: 0  
DB: 7 Gaps: 0

US-09-671-687A-3 (1-949) x US-10-972-079-20539 (1-600)

QY 714 LysAsnGluLysValGlyValProThrIleGlnLeuLeuGluTrpSerPheIleAsn 733  
|||||  
QY 714 LysAsnGluLysValGlyValProThrIleGlnLeuLeuGluTrpSerPheIleAsn 733  
|||||  
Db 266 AAAAAAGAAAGTTGGAGTCCGACAATYCACGAGTTACTGGAGTGGTCTTCATCAAC 207  
|||||

Db 329 AAAAAATGAGAAAGTTGGAGTCCGACAATYCACGAGTTACTGGAGTGGTCTTCATCAAC 270  
QY 734 SerAsnLeuLysPheAlaGlu 740  
Db 269 AGCAACTTGAAGTTTGCAGAG 249  
|||||

RESULT 6  
US-10-972-079-20540/c  
Sequence 20540, Application US/10972079  
GENERAL INFORMATION:  
APPLICANT: MMI GENOMICS, INC.  
APPLICANT: DENISE, Sue K.  
APPLICANT: ROSENFELD, David  
APPLICANT: KERR, Richard  
APPLICANT: BATES, Stephen  
APPLICANT: HOLM, Tom  
TITLE OF INVENTION: METHODS & SYSTEMS FOR INFERRING TRAITS TO BREED & MANAGE NON-BEER  
TITLE OF INVENTION: LIVESTOCK  
FILE REFERENCE: MM1110-2  
CURRENT APPLICATION NUMBER: US/10/972,079  
CURRENT FILING DATE: 2004-10-22  
PRIOR APPLICATION NUMBER: US 60/514,333  
PRIOR FILING DATE: 2003-10-24  
NUMBER OF SEQ ID NOS: 96631  
SOFTWARE: Patentin version 3.1  
SEQ ID NO 20540  
LENGTH: 600  
TYPE: DNA  
ORGANISM: Chicken 19866894231145\_3  
US-10-972-079-20540

Alignment Scores:  
Pred. No.: 1.15e-16 Length: 600  
Score: 27.00 Matches: 27  
Percent Similarity: 100.00% Conservatives: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 2.85% Indels: 0  
DB: 7 Gaps: 0

US-09-671-687A-3 (1-949) x US-10-972-079-20540 (1-600)

QY 714 LysAsnGluLysValGlyValProThrIleGlnLeuLeuGluTrpSerPheIleAsn 733  
|||||  
Db 266 AAAAAAGAAAGTTGGAGTCCGACAATYCACGAGTTACTGGAGTGGTCTTCATCAAC 207  
|||||

QY 734 SerAsnLeuLysPheAlaGlu 740  
Db 206 AGCAACTTGAAGTTTGCAGAG 186  
|||||

RESULT 7  
US-10-972-079-19097  
Sequence 19097, Application US/10972079  
GENERAL INFORMATION:  
APPLICANT: MMI GENOMICS, INC.  
APPLICANT: DENISE, Sue K.  
APPLICANT: ROSENFELD, David  
APPLICANT: KERR, Richard  
APPLICANT: BATES, Stephen  
APPLICANT: HOLM, Tom  
TITLE OF INVENTION: METHODS & SYSTEMS FOR INFERRING TRAITS TO BREED & MANAGE NON-BEER  
TITLE OF INVENTION: LIVESTOCK  
FILE REFERENCE: MM1110-2  
CURRENT APPLICATION NUMBER: US/10/972,079  
CURRENT FILING DATE: 2004-10-22  
PRIOR APPLICATION NUMBER: US 60/514,333  
PRIOR FILING DATE: 2003-10-24  
NUMBER OF SEQ ID NOS: 96631  
SOFTWARE: Patentin version 3.1  
SEQ ID NO 19097  
LENGTH: 600  
TYPE: DNA  
ORGANISM: Chicken 19866894228522\_1  
US-10-972-079-19097

Alignment Scores:  
Pred. No.: 154 Length: 600  
Score: 9.00 Matches: 9  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 0.95% Indels: 0  
DB: 7 Gaps: 0

US-09-671-687A-3 (1-949) x US-10-972-079-19097 (1-600)

Qy 601 PheCysLeuPheAlaPheSerSerVal 609  
Db 178 TTCGTGTTGTTGCTTTTCTTCAGTC 204

## RESULT 8

US-60-655-875-114267  
; Sequence 114267, Application US/60655875  
; GENERAL INFORMATION:  
; APPLICANT: Boukharov, Andrey  
; APPLICANT: Du, Zijiang  
; APPLICANT: Guo, Liang  
; APPLICANT: Kovalic, David  
; APPLICANT: Lu, Maolong  
; APPLICANT: McCarter, James  
; APPLICANT: Miller, Nancy  
; APPLICANT: Williams, Deryck  
; APPLICANT: Vaudin, Mark  
; APPLICANT: Wu, Wei  
; TITLE OF INVENTION: METHODS FOR GENETIC CONTROL OF HETERODERA INFESTATIONS  
; FILE REFERENCE: 38-21(53885)  
; CURRENT APPLICATION NUMBER: US/60/655, 875  
; CURRENT FILING DATE: 2005-02-24  
; NUMBER OF SEQ ID NOS: 171306  
; SEQ ID NO 114267  
; LENGTH: 1003  
; TYPE: DNA  
; ORGANISM: Heterodera glycines  
; FEATURE:  
; OTHER INFORMATION: Promoter: Source sequence=SeqID\_17891; Location=516..1518; Strand  
; OTHER INFORMATION: Gene=SeqID\_82441  
US-60-655-875-114267

Alignment Scores:  
Pred. No.: 246 Length: 1003  
Score: 9.00 Matches: 9  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 0.95% Indels: 0  
DB: 12 Gaps: 0

US-09-671-687A-3 (1-949) x US-60-655-875-114267 (1-1003)

Qy 601 PheCysLeuPheAlaPheSerSerVal 609  
Db 902 TTCGTGTTGTTGCTTTTCTTCAGTT 928

## RESULT 9

US-60-655-875-75550/c  
; Sequence 75550, Application US/60655875  
; GENERAL INFORMATION:  
; APPLICANT: Boukharov, Andrey  
; APPLICANT: Du, Zijiang  
; APPLICANT: Guo, Liang  
; APPLICANT: Kovalic, David  
; APPLICANT: Lu, Maolong  
; APPLICANT: McCarter, James  
; APPLICANT: Miller, Nancy  
; APPLICANT: Williams, Deryck  
; APPLICANT: Vaudin, Mark  
; APPLICANT: Wu, Wei  
; TITLE OF INVENTION: METHODS FOR GENETIC CONTROL OF HETERODERA INFESTATIONS

; TITLE OF INVENTION: IN PLANTS AND COMPOSITIONS THEREOF  
; FILE REFERENCE: 38-21(53885)  
; CURRENT APPLICATION NUMBER: US/60/655, 875  
; CURRENT FILING DATE: 2005-02-24  
; NUMBER OF SEQ ID NOS: 171306  
; SEQ ID NO 75550  
; LENGTH: 1854  
; TYPE: DNA  
; ORGANISM: Heterodera glycines  
; FEATURE:  
; OTHER INFORMATION: Genomic contig location: Contig\_ID=SeqID\_1731; Location=4308..616  
; OTHER INFORMATION: Strand=+

; FEATURE:  
; OTHER INFORMATION: SCN-specific regions: Sequence locations=1-164,211-538,560-629,65  
; OTHER INFORMATION: -758,780-971,993-1250,1378-1395,1454-1577,1599-1746,1771-1854  
; FEATURE:  
; OTHER INFORMATION: Gene Ontology: Mol. function=heat shock protein activity; Cellula  
; OTHER INFORMATION: component=cytosol; Biological process=protein folding  
US-60-655-875-75550

Alignment Scores:  
Pred. No.: 433 Length: 1854  
Score: 9.00 Matches: 9  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 0.95% Indels: 0  
DB: 12 Gaps: 0

US-09-671-687A-3 (1-949) x US-60-655-875-75550 (1-1854)

Qy 730 SerPheIleAsnSerAsnLeuLysPhe 738  
Db 1433 TCATTTATTAATCAATCTTAATTT 1407

## RESULT 10

US-60-655-875-17891  
; Sequence 17891, Application US/60655875  
; GENERAL INFORMATION:  
; APPLICANT: Boukharov, Andrey  
; APPLICANT: Du, Zijiang  
; APPLICANT: Guo, Liang  
; APPLICANT: Kovalic, David  
; APPLICANT: Lu, Maolong  
; APPLICANT: McCarter, James  
; APPLICANT: Miller, Nancy  
; APPLICANT: Williams, Deryck  
; APPLICANT: Vaudin, Mark  
; APPLICANT: Wu, Wei  
; TITLE OF INVENTION: METHODS FOR GENETIC CONTROL OF HETERODERA INFESTATIONS  
; FILE REFERENCE: 38-21(53885)  
; CURRENT APPLICATION NUMBER: US/60/655, 875  
; CURRENT FILING DATE: 2005-02-24  
; NUMBER OF SEQ ID NOS: 171306  
; SEQ ID NO 17891  
; LENGTH: 2900  
; TYPE: DNA  
; ORGANISM: Heterodera glycines  
US-60-655-875-17891

Alignment Scores:  
Pred. No.: 654 Length: 2900  
Score: 9.00 Matches: 9  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 0.95% Indels: 0  
DB: 12 Gaps: 0

US-09-671-687A-3 (1-949) x US-60-655-875-17891 (1-2900)

Qy 601 PheCysLeuPheAlaPheSerSerVal 609  
Db 1417 TTCGTGTTGTTGCTTTTCTTCAGTT 1443



```
RESULT 11
US-60-655-875-58260/c
; Sequence 58260, Application US/60655875
; GENERAL INFORMATION:
; APPLICANT: Boukharov, Andrey
; APPLICANT: Du, Zijing
; APPLICANT: Guo, Liang
; APPLICANT: Kovalic, David
; APPLICANT: Lu, Maolong
; APPLICANT: McCarter, James
; APPLICANT: Miller, Nancy
; APPLICANT: Williams, Deryck
; APPLICANT: Vaudin, Mark
; APPLICANT: Wu, Wei
; TITLE OF INVENTION: METHODS FOR GENETIC CONTROL OF HETERODERA INFESTATIONS
; TITLE OF INVENTION: IN PLANTS AND COMPOSITIONS THEREOF
; FILE REFERENCE: 38-21(53885)
; CURRENT APPLICATION NUMBER: US/60/655,875
; CURRENT FILING DATE: 2005-02-24
; NUMBER OF SEQ ID NOS: 171306
; SEQ ID NO 58260
; LENGTH: 3949
; TYPE: DNA
; ORGANISM: Heterodera glycines
; FEATURE:
; OTHER INFORMATION: Genomic contig location: Contig_ID=SeqID_5500; Location=408...4356
; OTHER INFORMATION: Strand=+
; FEATURE:
; OTHER INFORMATION: SCN-specific regions: Sequence locations=1-298,343-353,376-457,48
; OTHER INFORMATION: -804,826-991,1013-1382,1404-1630,1662-1822,1844-1871,1893-2224,22
; OTHER INFORMATION: -2315,2337-2458,2483-2578,2600-2758,2780-2942,2964-3670,3693-3949
; FEATURE:
; OTHER INFORMATION: Gene Ontology: Mol. function=translation elongation factor activ
; OTHER INFORMATION: Cellular component=mitochondrial ribosome; Biological process
; OTHER INFORMATION: =gametogenesis
US-60-655-875-58260

Alignment Scores:
Pred. No.: 868 Length: 3949
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.95% Indels: 0
DB: 12 Gaps: 0

US-09-671-687A-3 (1-949) x US-60-655-875-58260 (1-3949)

Qy 391 SerSerProLeuGlnProPro 399
Db 2933 TCCTCCCTCCTTTACAGCGGCCCT 2907

RESULT 12
US-60-655-875-1731/c
; Sequence 1731, Application US/60655875
; GENERAL INFORMATION:
; APPLICANT: Boukharov, Andrey
; APPLICANT: Du, Zijing
; APPLICANT: Guo, Liang
; APPLICANT: Kovalic, David
; APPLICANT: Lu, Maolong
; APPLICANT: McCarter, James
; APPLICANT: Miller, Nancy
; APPLICANT: Williams, Deryck
; APPLICANT: Vaudin, Mark
; APPLICANT: Wu, Wei
; TITLE OF INVENTION: METHODS FOR GENETIC CONTROL OF HETERODERA INFESTATIONS
; TITLE OF INVENTION: IN PLANTS AND COMPOSITIONS THEREOF
; FILE REFERENCE: 38-21(53885)
; CURRENT APPLICATION NUMBER: US/60/655,875
; CURRENT FILING DATE: 2005-02-24
; NUMBER OF SEQ ID NOS: 171306
; SEQ ID NO 1731
```

```
; LENGTH: 6161
; TYPE: DNA
; ORGANISM: Heterodera glycines
US-60-655-875-1731

Alignment Scores:
Pred. No.: 1.31e+03 Length: 6161
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.95% Indels: 0
DB: 12 Gaps: 0

US-09-671-687A-3 (1-949) x US-60-655-875-1731 (1-6161)

Qy 730 SerPheIleAsnSerAsnLeuLysPhe 738
Db 5740 TCATTTATTAAATTCAAATCTTAATTT 5714

RESULT 13
US-60-655-875-5500/c
; Sequence 5500, Application US/60655875
; GENERAL INFORMATION:
; APPLICANT: Boukharov, Andrey
; APPLICANT: Du, Zijing
; APPLICANT: Guo, Liang
; APPLICANT: Kovalic, David
; APPLICANT: Lu, Maolong
; APPLICANT: McCarter, James
; APPLICANT: Miller, Nancy
; APPLICANT: Williams, Deryck
; APPLICANT: Vaudin, Mark
; APPLICANT: Wu, Wei
; TITLE OF INVENTION: METHODS FOR GENETIC CONTROL OF HETERODERA INFESTATIONS
; TITLE OF INVENTION: IN PLANTS AND COMPOSITIONS THEREOF
; FILE REFERENCE: 38-21(53885)
; CURRENT APPLICATION NUMBER: US/60/655,875
; CURRENT FILING DATE: 2005-02-24
; NUMBER OF SEQ ID NOS: 171306
; SEQ ID NO 5500
; LENGTH: 6758
; TYPE: DNA
; ORGANISM: Heterodera glycines
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (6672)..(6672)
; OTHER INFORMATION: n is a, c, g, or t
US-60-655-875-5500

Alignment Scores:
Pred. No.: 1.42e+03 Length: 6758
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.95% Indels: 0
DB: 12 Gaps: 0

US-09-671-687A-3 (1-949) x US-60-655-875-5500 (1-6758)

Qy 391 SerSerProLeuGlnProPro 399
Db 3340 TCCTCCCTCCTTTACAGCGGCCCT 3314

RESULT 14
US-10-287-436A-956
; Sequence 956, Application US/10287436A
; GENERAL INFORMATION:
; APPLICANT: CHILDREN'S HOSPITAL MEDICAL CENTER
; TITLE OF INVENTION: METHOD FOR DIAGNOSIS AND TREATMENT OF
; TITLE OF INVENTION: RHEUMATOID ARTHRITIS
; FILE REFERENCE: 10872.514696
; CURRENT APPLICATION NUMBER: US/10/287,436A
; CURRENT FILING DATE: 2002-10-31
```

```
; PRIOR APPLICATION NUMBER: US 60/336,220
; PRIOR FILING DATE: 2001-10-31
; NUMBER OF SEQ ID NOS: 1446
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 956
; LENGTH: 149034
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-287-436A-956
```

```
Alignment Scores:
Pred. No.:      2.44e+04      Length:      149034
Score:          9.00          Matches:      9
Percent Similarity: 100.00%    Conservative: 0
Best Local Similarity: 100.00% Mismatches:      0
Query Match:      0.95%       Indels:         0
DB:              7           Gaps:           0
```

US-09-671-687A-3 (1-949) x US-10-287-436A-956 (1-149034)

```
Qy      521 LeuLysLysAlaLeuPheValLysLeu 529
      |||||
Db      97267 CTTAAAGGCGATTTTGTGAACCTT 97293
```

# RESULT 15

```
US-11-036-317-90968
; Sequence 90968, Application US/11036317
; GENERAL INFORMATION:
; APPLICANT: Williams, Alan
; APPLICANT: Blume, John
; TITLE OF INVENTION: Method of Analysis of Alternative Splicing in Mouse
; FILE REFERENCE: 3654.1
; CURRENT APPLICATION NUMBER: US/11/036,317
; CURRENT FILING DATE: 2005-01-13
; PRIOR APPLICATION NUMBER: US 60/536,639
; PRIOR FILING DATE: 2004-01-13
; NUMBER OF SEQ ID NOS: 991174
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 90968
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Mus musculus
US-11-036-317-90968
```

```
Alignment Scores:
Pred. No.:      84.2          Length:      25
Score:          8.00          Matches:      8
Percent Similarity: 100.00%    Conservative: 0
Best Local Similarity: 100.00% Mismatches:      0
Query Match:      0.84%       Indels:         0
DB:              10          Gaps:           0
```

US-09-671-687A-3 (1-949) x US-11-036-317-90968 (1-25)

```
Qy      142 ProGlyValValArgPheArgGly 149
      |||||
Db      1 CCAGGAGTTGTACGCTTCAGAGGA 24
```

Search completed: April 17, 2005, 13:54:51  
Job time : 719 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: April 17, 2005, 06:40:33 ; Search time 377 Seconds  
(without alignments)

4118.903 Million cell updates/sec

Title: US-09-671-687A-3

Perfect score: 949

Sequence: 1 MSSGLWSQKVTSPYWEERI.....RLLCDAYMCMYQSTMSLYK 949

Scoring table:

OLIGO  
Xgapop 60.0 , Xgapext 60.0  
Ygapop 60.0 , Ygapext 60.0  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 1202784 seqs, 818138359 residues

Word size: 1

Total number of hits satisfying chosen parameters: 2400006

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp  
-Q=/cgn2\_1/USPTO.spool\_p/US09671687/runat\_15042005\_163039\_11977/app\_query.fasta\_1.1095  
-DB=Issued Patents NA -QMT=fastap -SUFFIX=Oligo.rni -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09671687@cgn\_1\_105 @runat\_15042005\_163039\_11977 -NCPU=6 -ICPU=3  
-NO MMAP -LARGESQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

- Issued Patents\_NA:\*
- 1: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq:\*
  - 2: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq:\*
  - 3: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq:\*
  - 4: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq:\*
  - 5: /cgn2\_6/ptodata/1/ina/PCTUS\_COMB.seq:\*
  - 6: /cgn2\_6/ptodata/1/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	641	67.5	2523	4	US-09-620-312D-290	Sequence 290, Appl
2	165	17.4	2116	4	US-09-646-403-3	Sequence 3, Appl1
3	36	3.8	110	4	US-09-513-999C-20085	Sequence 20085, A
C 4	9	0.9	747	3	US-08-998-416-246	Sequence 246, App
	5	9	6574	3	US-09-221-017B-1097	Sequence 1097, Ap
6	9	0.9	20284	4	US-09-526-193A-21	Sequence 21, Appl
C 7	9	0.9	53806	4	US-09-949-016-12572	Sequence 12572, A
C 8	8	0.8	255	3	US-09-134-001C-2194	Sequence 2194, Ap
C 9	8	0.8	266	4	US-09-232-785-240	Sequence 240, App
10	8	0.8	601	4	US-09-949-016-52259	Sequence 52259, A
C 11	8	0.8	601	4	US-09-949-016-52319	Sequence 52319, A
C 12	8	0.8	601	4	US-09-949-016-52320	Sequence 52320, A

C 13	8	0.8	601	4	US-09-949-016-52321 Sequence 52321, A
C 14	8	0.8	601	4	US-09-949-016-56429 Sequence 56429, A
C 15	8	0.8	601	4	US-09-949-016-56450 Sequence 56450, A
C 16	8	0.8	601	4	US-09-949-016-56471 Sequence 56471, A
C 17	8	0.8	601	4	US-09-949-016-96111 Sequence 96111, A
C 18	8	0.8	601	4	US-09-949-016-192243 Sequence 192243, A
19	8	0.8	601	4	US-09-949-016-195831 Sequence 195831, A
20	8	0.8	601	4	US-09-949-016-195832 Sequence 195832, A
21	8	0.8	687	4	US-09-252-991A-5646 Sequence 5646, Ap
22	8	0.8	702	4	US-09-252-991A-5549 Sequence 5549, Ap
23	8	0.8	771	4	US-09-902-540-3210 Sequence 3210, Ap
24	8	0.8	1030	4	US-09-621-976-18958 Sequence 18958, A
25	8	0.8	1098	4	US-09-270-767-10775 Sequence 10775, A
C 26	8	0.8	1164	4	US-09-107-532A-2278 Sequence 2278, Ap
27	8	0.8	1416	4	US-09-583-110-251 Sequence 251, App
28	8	0.8	1428	4	US-09-107-433-2425 Sequence 2425, Ap
29	8	0.8	1713	4	US-09-252-991A-3439 Sequence 3439, Ap
C 30	8	0.8	1938	4	US-09-252-991A-5524 Sequence 5524, Ap
C 31	8	0.8	2112	4	US-09-252-991A-3494 Sequence 3494, Ap
32	8	0.8	2127	4	US-09-252-991A-3450 Sequence 3450, Ap
33	8	0.8	2187	4	US-09-252-991A-5616 Sequence 5616, Ap
34	8	0.8	2481	4	US-09-902-540-5228 Sequence 5228, Ap
C 35	8	0.8	3719	4	US-08-956-171B-344 Sequence 344, App
C 36	8	0.8	3719	4	US-08-781-986A-344 Sequence 344, App
37	8	0.8	5107	4	US-09-358-383C-15 Sequence 15, Appl
38	8	0.8	5414	4	US-09-949-016-14322 Sequence 14322, A
C 39	8	0.8	5591	4	US-09-949-016-12044 Sequence 12044, A
C 40	8	0.8	5592	4	US-09-949-016-16387 Sequence 16387, A
41	8	0.8	5955	4	US-09-358-383C-14 Sequence 14, Appl
42	8	0.8	6651	4	US-09-302-540-4944 Sequence 4944, Ap
43	8	0.8	7515	4	US-09-328-352-1695 Sequence 1695, Ap
44	8	0.8	7886	3	US-09-453-702B-102 Sequence 102, App
45	8	0.8	8729	3	US-09-453-702B-258 Sequence 258, App

ALIGNMENTS

RESULT 1

US-09-620-312D-290  
; Sequence 290, Application US/09620312D  
; Patent No. 6569662  
; GENERAL INFORMATION:  
; APPLICANT: Tang, Y. Tom  
; APPLICANT: Liu, Chenghua  
; APPLICANT: Asundi, Vinod  
; APPLICANT: Zhang, Jie  
; APPLICANT: Ren, Feiyan  
; APPLICANT: Chen, Rui-hong  
; APPLICANT: Zhao, Qing A.  
; APPLICANT: Wehrman, Tom  
; APPLICANT: Xue, Aidong J.  
; APPLICANT: Yang, Yonghong  
; APPLICANT: Wang, Jian-Rui  
; APPLICANT: Zhou, Ping  
; APPLICANT: Ma, Yunding  
; APPLICANT: Wang, Dunrui  
; APPLICANT: Wang, Zhiwei  
; APPLICANT: John Tillinghast  
; APPLICANT: Drmanac, Radoje T.  
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and  
; TITLE OF INVENTION: Polypeptides  
; FILE REFERENCE: 784CIP2B  
; CURRENT APPLICATION NUMBER: US/09/620,312D  
; CURRENT FILING DATE: 2000-07-19  
; PRIOR APPLICATION NUMBER: 09/552,317  
; PRIOR FILING DATE: 2000-04-25  
; PRIOR APPLICATION NUMBER: 09/488,725  
; PRIOR FILING DATE: 2000-01-21  
; NUMBER OF SEQ ID NOS: 1105  
; SOFTWARE: pt\_FL\_genes Version 1.0  
; SEQ ID NO 290  
; LENGTH: 2523  
; TYPE: DNA

ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (81)..(2138)  
US-09-620-312D-290

## Alignment Scores:

Pred. NO.: 0 Length: 2523  
Score: 641.00 Matches: 668  
Percent Similarity: 99.55% Conservativity: 0  
Best Local Similarity: 99.55% Mismatches: 0  
Query Match: 67.54% Indels: 3  
DB: 4 Gaps: 0

US-09-671-687A-3 (1-949) x US-09-620-312D-290 (1-2523)

QY 282 LeuCysSerPheAlaCysValGluSerThrIleLeuLeuHisIleAsnAspIlePro 301  
DB CTTTGTAGTTTGGCGTGTGAAAGTACAATTTCTATTGGCATCATCAATGATATATCCCA 188  
QY 302 GluSerValThrGlnGluArgProProLysLeuAlaPheMetSerArgGlyValGly 321  
DB GAGAGTGTGACGAGAAAGGCGCTCCCAACTTGGCTTTATGTCAGAGGTGTGG 248  
QY 322 AspLysGlySerSerHisAsnLysProLysAlaThrGlySerThrSerAspProGly 341  
DB GACAAAGGTTTCATCCAGTCATAATAAACCAAGGCTACAGGATCTACCTCAGACCTGGA 308  
QY 342 AsnArg---ArgSerGluLeuPheThrLeuAsnGlySerSerValAspSerGlnPro 360  
DB AATAGAACACAGATCTGAATTTATATACCTTAAATGGGTCTCTGTTCAGCTCAACCA 368  
QY 361 GlnSerLysSerLysAsnThrTyrIleAspGluValAlaGluAspProAlaLysSer 380  
DB CAATCCAAATCAAAAATACATGTGATGATGAAGTTGCAAGACCTCGCAAAATCT 428  
QY 381 LeuThrGluIleSerThrAspPheAspArgSerSerProLeuGlnProProVal 400  
DB CTTACAGAGATATCTACAGACTTTGACCGTCTTCCACCACCATCCAGCTCTCTCTGTG 488  
QY 401 AsnSerLeuThrThrGluAsnArgPheHisSerLeuProPheSerLeuThrLysMetPro 420  
DB AACTCCTACGTGACCGGAGACAGATTCACCTCTTTACCATTTCAGTCTCACCAGAGTGCC 548  
QY 421 AsnThrAsnGlySerIleGlyHisSerProLeuSerLeuSerAlaGlnSerValMetGlu 440  
DB AATACCAATGGAAGTATTTGGCCACAGTCCACTTTCTCTGTCAGCCAGCTGTGTAATGGA 608  
QY 441 GluLeuAsnThrAlaProValGlnGluSerProProLeuAlaMetProProGlyAsnSer 460  
DB GAGCTAAACACTGCACCCGTCCAGAGAGTCCACCTTGGCCATGCTCTCTGGGAACTCA 668  
QY 461 HisGlyLeuGluValGlySerLeuAlaGluValLysGluAsnProProPheThrGlyVal 480  
DB CATGCTCTAGAAGTGGGCTCATTTGGCTGAAAGTTAAGGAAACCCCTCTTTCTATGGGTA 728  
QY 481 IleArgTrpIleGlyGlnProProGlyLeuAsnGluValLeuAlaGlyLeuGluLeuGlu 500  
DB ATCCGTTGGATCGGTCAGCCACGAGCTGAATGAAGTGGCTCGCTGGACTGGAATCGGAA 788  
QY 501 AspGluCysAlaGlyCysThrAspGlyThrPheArgGlyThrArgTyrPheThrCysAla 520  
DB GATGAGTGTGAGGCTGTACGGATGGAACCTTCAGAGGCATCTCGGTATTTTACCTGTGCC 848  
QY 521 LeuLysLysAlaLeuPheValLysLeuLysSerCysArgProAspSerArgPheAlaSer 540  
DB CTGAAGAAGCGCTGTTTGTGAAACTGAAGAGCTGCAGGCTGACTCTAGGTTTCATCA 908  
QY 541 LeuGlnProValSerAsnGlnIleGluArgCysAsnSerLeuAlaPheGlyGlyTyrLeu 560  
DB TTGACGCGGTTTCCAAATCAGATTCAGGCGCTGTAATCTTTTAGCATTTGGAGGCTACTTA 968  
QY 561 SerGluValGluGluAsnThrProProLysMetGluLysGluGlyLeuGluIleMet 580

DB AGTGAAGTAGTAGAAGAAAATACTCCACCAAAAATGGAAAAGAGGCTTGGAGATAATG 1028  
QY IleGlyLysValLysGlyIleGlnGlyHisIleValSerCysTyrLeuAspSerThrLeu 600  
DB ATTGGGAAGAAGAAAGGCATCCAGGTCATTACAATTTCTTTGATCTTAGACTCAACCTTA 1088  
QY PheCysLeuPheAlaPheSerSerValLeuAspThrValLeuLeuArgProLysGluLys 620  
DB TTCTGCTTATTTGCTTTTGTAGTTCTGTTCTGGACACTGTGTACTTAGACCCCAAGAAAG 1148  
QY AsnAspValGluTyrTyrSerGluThrGlnGluLeuLeuArgThrGluIleValAsnPro 640  
DB AACGATGTAGAATAATTATAGTGAACCAACCAAGAGCTACTCAGACACAGAAATTTTAATCT 1208  
QY LeuArgIleTyrGlyTyrValCysAlaThrLysIleMetLysLeuArgLysIleLeuGlu 660  
DB CTGAGATATATGGAATATGTGTGTGCCACAAAATTTATGAACCTGAGGAAATATCTTGA 1268  
QY LysValGluAlaAlaSerGlyPheThrSerGluLysAspProGluGluPheLeuAsn 680  
DB AAGGTGAGGCTGCATCAGGATTTACCTCTGAAGAAAAGATCTCTGAGGAATTTCTTGAAT 1328  
QY IleLeuPheHisIleLeuArgValGluProLeuLeuLysIleArgSerAlaGlyGln 700  
DB ATTCTGTTTCATCATATTTAAGGTTAGAACCTTTGCTAAAAATAAGATCAGCAGTCAA 1388  
QY LysValGlnAspCysTyrPheTyrGlnIlePheMetGluLysAsnGluLysValGlyVal 720  
DB AAGGTACAGATTTGTTACTTCTATCAAAATTTTATGGAAAAAATGAGAAAGTTGGCGTT 1448  
QY ProThrIleGlnGlnLeuLeuGluTyrPheIleAsnSerAsnLysPheAlaGlu 740  
DB CCCCAATTCAGCAGTGTGTAGATGTTCTTTTATCAACAGCTAACTGAAATTTGCAGAG 1508  
QY AlaProSerCysLeuIleIleGlnMetProArgPheGlyLysAspPheLysLeuPheLys 760  
DB GCACCATCATGTCGATTATTAGATGCTCGATTTGGAAAAAGACTTTAAACATAATTTAAA 1568  
QY LysIlePheProSerLeuGluLeuAsnIleThrAspLeuLeuGluAspThrProArgGln 780  
DB AAAATTTTCTCTCTCTGGAATTTAATAATAACAGATTTTACTTGAAGACACTCCAGACAG 1628  
QY CysArgIleCysGlyValLeuAlaMetTyrGluCysArgGluCysTyrAspAspProAsp 800  
DB TGCCGATATGTGGAGGCTTGCATGATGATGATGATGATGATGATGATGATGATGATGATG 1688  
QY IleSerAlaGlyLysIleLysGlnPheCysLysThrCysAsnThrGlnValHisLeuHis 820  
DB ATCTCAGCTGGAAAATCAAGCAGTTTGTAAAACTTCACACACTCAAGTCCACCTTCAT 1748  
QY ProLysArgLeuAsnHisLysTyrAsnProValSerLeuProLysAspLeuProAspTrp 840  
DB CCGAAGAGGCTGAATCATATAATATAACCCAGTGTCTCTCCCAAGACTTACCCGACTGG 1808  
QY AspTrpArgHisGlyCysIleProCysGlnAsnMetGluLeuPheAlaValLeuCysIle 860  
DB GACTGGAGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1868  
QY GluThrSerHisTyrValAlaPheValLysThrGlyLysAspAspSerAlaTrpLeuPhe 880  
DB GAAACAGCCACTATGTTGCTTTGTGAATATGGAGAGGAGGATCTTGCCTGGCTCTTC 1928  
QY PheAspSerMetAlaAspArgAspGlyGlnAsnGlyPheAsnIleProGlnValThr 900  
DB TTTGACAGACATGGCCGATCGGATCGTGTGATGATGATGATGATGATGATGATGATGATG 1988  
QY ProCysProGluValGlyGluTyrLeuLysMetSerLeuGluAspLeuHisSerLeuAsp 920  
DB CCATGCCCAAGTAGGAGAGTACTTTGAAGATGCTCTCTGGAAGACCTGCTATCTTCTTGGAC 2048  
QY SerArgArgIleGlnGlyCysAlaArgGluLeuCysAspAlaTyrMetCysMetTyr 940

Db 2049 TCCAGGAGATCCCAAGGCTGTCACGAAGACTGCTTTGTGATGCATATATGTGCATGTAC 2108  
Qy 941 GlnSerProThrMetSerLeuTyrLys 949  
Db 2109 CAGAGTCCAAACAATGAGTTTGTACAA 2135  
RESULT 2  
US-09-646-403-3  
; Sequence 3, Application US/09646403  
; Patent No. 6734174  
; GENERAL INFORMATION:  
; APPLICANT: WALLACH, David  
; APPLICANT: KOVALENKO, Andrei  
; TITLE OF INVENTION: MODULATORS OF THE FUNCTION OF RECEPTORS OF THE TNF/NGF RECEPTOR F  
; TITLE OF INVENTION: AND OTHER PROTEINS  
; FILE REFERENCE: WALLACH-27  
; CURRENT APPLICATION NUMBER: US/09/646,403  
; CURRENT FILING DATE: 2000-09-18  
; PRIOR FILING DATE: IL 123758  
; PRIOR APPLICATION NUMBER: PCT/IL99/00158  
; PRIOR FILING DATE: 1998-03-19  
; PRIOR APPLICATION NUMBER: IL 126024  
; PRIOR FILING DATE: 1998-09-01  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 3  
; LENGTH: 2116  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (691)..(691)  
; OTHER INFORMATION: n is unknown.  
US-09-646-403-3  
Alignment Scores:  
Pred. No.: 1.38e-164 Length: 2116  
Score: 165.00 Matches: 178  
Percent Similarity: 99.44% Conservative: 0  
Best Local Similarity: 99.44% Mismatches: 0  
Query Match: 17.39% Indels: 1  
DB: 4 Gaps: 0  
US-09-671-687A-3 (1-949) x US-09-646-403-3 (1-2116)  
Qy 387 AspPheAspArgSerSerProProLeuGlnProProProValAsnSerLeuThrThrGlu 406  
Db 15 GACTTTGACCCGTTCTTCCACCACCACTCCAGCCTCTCCTGTGNACTCACTGACCACCGAG 74  
Qy 407 AsnArgPheHisSerLeuProPheSerLeuThrLysMetProAsnThrAsnGlySerIle 426  
Db 75 AACAGATTCCACTCTTTACCATTCACTGCTCACCAGATGCCCAATACCAATGGAAGTATT 134  
Qy 427 GlyHisSerProLeuSerLeuSerAlaGlnSerValMetGluLeuAsnThrAlaPro 446  
Db 135 GGCCACAGTCACCTTTCTGTGCAGCCAGCTGTGAATGGAAGAGCTAAACACTGCACCC 194  
Qy 447 ValGlnGluSerProProLeuAlaMetProProGlyAsnSerHisGlyLeuGluValGly 466  
Db 195 GTCCAGAGAGTCCACCTTTGGCCATGCTCTCGGNACTCACTGCTAGAGTGGGC 254  
Qy 467 SerLeuAlaGluValLysGluAsnProProPheTyrGlyValIleArgTrpIleGlyGln 486  
Db 255 TCAITGGCTGAAGTTAAGGAGAACCTCTCTTCTATGGGGTAATCGTTGGATCGGTCAG 314  
Qy 487 ProProGlyLeuAsnGluValLeuAlaGlyLeuGluLeuAspGluCysAlaGlyCys 506  
Db 315 CCACAGAGACTGAATGAAGTGTCTCGCTGAGCTGGAAGTGAAGATGAGTGCAGGCTGT 374  
Qy 507 ThrAspGlyThrPheArgGlyThrArgTyrPheThrCysAlaLeuLysLysAlaLeuPhe 526  
Db 375 ACGATGGAACTTCAGAGGCACTCGGTATTTCATCTGTGCCCTGNAGAGGCGGTGTTT 434

Qy 527 ValLysLeuLysSerCysArgProAspSerArgPheAlaSerLeuGlnProValSerAsn 546  
Db 435 GTGAAACTGAAGAGCTGCAGGCTGACTCTAGGTTTGCATCATTTGCAGCGGTTTCCAAT 494  
Qy 547 Gln-IleGluArgCysAsnSerLeuAlaPheGlyGlyTyrLeuSerGluValVal 564  
Db 495 CAAGATTGAGCGCTGTAACCTCTTTAGCATTTGGAGGCTACTTTAAGTGAAGTAGTG 549

## RESULT 3

US-09-513-999C-20085  
; Sequence 20085, Application US/09513999C  
; Patent No. 6783961  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Duclert, A.  
; APPLICANT: Giordano, J.Y.  
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.  
; FILE REFERENCE: 59.US2.REG  
; CURRENT APPLICATION NUMBER: US/09/513,999C  
; CURRENT FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/122,487  
; PRIOR FILING DATE: 1999-02-26  
; NUMBER OF SEQ ID NOS: 36681  
; SOFTWARE: Patent.pm  
; SEQ ID NO 20085  
; LENGTH: 110  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-513-999C-20085

Alignment Scores:  
Pred. No.: 4.75e-29 Length: 110  
Score: 36.00 Matches: 36  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 3.79% Indels: 0  
DB: 4 Gaps: 0

US-09-671-687A-3 (1-949) x US-09-513-999C-20085 (1-110)

Qy 595 TyrLeuAspSerThrLeuPheCysLeuPheAlaPheSerValLeuAspThrValLeu 614  
Db 1 TACTTAGACTCAACCTTATCTGCTTATTGCTTTAGTCTGTTCTGGACACTGTGTTA 60  
Qy 615 LeuArgProLysGluLysAsnAspValGluTyrTyrSerGluThrGln 630  
Db 61 CTTAGACCCCAAGAAAGAACGATGTAGATAATTATAGTGAAACCCAA 108

## RESULT 4

US-08-998-416-246/c  
; Sequence 246, Application US/08998416  
; Patent No. 6239264  
; GENERAL INFORMATION:  
; APPLICANT: Philippsen, Peter  
; APPLICANT: Pohlmann, Rainer  
; APPLICANT: Steiner, Sabine  
; APPLICANT: Mohr, Christine  
; APPLICANT: Wendland, Jurgен  
; APPLICANT: Knechtle, Philipp  
; APPLICANT: Rebschlung, Corinne  
; TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSSYPHII  
; TITLE OF INVENTION: AND USES THEREOF  
; NUMBER OF SEQUENCES: 1152  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: No. 6239264artis Corporation  
; STREET: 3054 Cornwallis Road  
; CITY: Research Triangle Park  
; STATE: No. 6239264th Carolina  
; COUNTRY: USA  
; ZIP: 27709  
; COMPUTER READABLE FORM:

;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.30  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/998,416  
;; FILING DATE: 24-DEC-1997  
;; CLASSIFICATION: 435  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: CH 0016/97  
;; FILING DATE: 31-DEC-1996  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Meigs, J. Timothy  
;; REGISTRATION NUMBER: 38,241  
;; REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 919-541-8587  
;; TELEFAX: 919-541-8689  
;; INFORMATION FOR SEQ ID NO: 246:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 747 base pairs  
;; TYPE: nucleic acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: DNA (genomic)  
;; ORIGINAL SOURCE:  
;; ORGANISM: PAG1220RP  
US-08-998-416-246  
  
Alignment Scores:  
Pred. No.: 15 Length: 747  
Score: 9.00 Matches: 9  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 0.95% Indels: 0  
DB: 3 Gaps: 0  
  
US-09-671-687A-3 (1-949) x US-08-998-416-246 (1-747)  
  
QY 333 AlathrGlySerThrSerAspProGly 341  
Db 332 GCCACCGGGAGCACCAGTGATCCAGG 306  
  
RESULT 5  
US-09-221-017B-1097  
; Sequence 1097, Application US/09221017B  
; Patent No. 6444799  
; GENERAL INFORMATION:  
; APPLICANT: Ross, Bruce C.  
; TITLE OF INVENTION: P. GINGIVALIS NUCLEOTIDES AND USES THEREOF  
; NUMBER OF SEQUENCES: 1120  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: MORRISON & FOERSTER  
; STREET: 755 PAGE MILL ROAD  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304-1018  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: Windows  
; SOFTWARE: FastSeq for Windows Version 2.0b  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/221,017B  
; FILING DATE: 23-DEC-1998  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PP1182  
; FILING DATE: 31-DEC-1997  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PP1546  
; FILING DATE: 30-JAN-1998

;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: PP2911  
;; FILING DATE: 09-APR-1998  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: PCT/AU98/01023  
;; FILING DATE: 10-DEC-1998  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Monroy, Gladys H  
;; REGISTRATION NUMBER: 32,430  
;; REFERENCE/DOCKET NUMBER: 27340-20021.00  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 650-813-5600  
;; TELEFAX: 650-494-0792  
;; TELEX: 706141  
;; INFORMATION FOR SEQ ID NO: 1097:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 6574 base pairs  
;; TYPE: nucleic acid  
;; STRANDEDNESS: double  
;; TOPOLOGY: circular  
;; MOLECULE TYPE: DNA (genomic)  
;; HYPOTHETICAL: NO  
;; ANTI-SENSE: UNKNOWN  
;; ORIGINAL SOURCE:  
;; ORGANISM: PORYPHYROMONAS GINGIVALIS  
;; FEATURE:  
;; NAME/KEY: misc feature  
;; LOCATION: 1...6574  
US-09-221-017B-1097  
  
Alignment Scores:  
Pred. No.: 140 Length: 6574  
Score: 9.00 Matches: 9  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 0.95% Indels: 0  
DB: 3 Gaps: 0  
  
US-09-671-687A-3 (1-949) x US-09-221-017B-1097 (1-6574)  
  
QY 425 SerIleGlyHisSerProLeuSerLeu 433  
Db 208 TCGATAGGACACTCGCCATTGTCTTG 234  
  
RESULT 6  
US-09-526-193A-21  
; Sequence 21, Application US/09526193A  
; Patent No. 6617122  
; GENERAL INFORMATION:  
; APPLICANT: Hayden, Michael R.  
; APPLICANT: Brooks-Wilson, Angela R.  
; APPLICANT: Pimstone, Simon N.  
; TITLE OF INVENTION: METHODS AND REAGENTS FOR MODULATING  
; TITLE OF INVENTION: CHOLESTEROL LEVELS  
; FILE REFERENCE: 50110/002005  
; CURRENT APPLICATION NUMBER: US/09/526,193A  
; CURRENT FILING DATE: 2000-03-15  
; PRIOR APPLICATION NUMBER: 60/124,702  
; PRIOR FILING DATE: 1999-03-15  
; PRIOR APPLICATION NUMBER: 60/138,048  
; PRIOR FILING DATE: 1999-06-08  
; PRIOR APPLICATION NUMBER: 60/139,600  
; PRIOR FILING DATE: 1999-06-17  
; PRIOR APPLICATION NUMBER: 60/151,977  
; PRIOR FILING DATE: 1999-09-01  
; NUMBER OF SEQ ID NOS: 287  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 21  
; LENGTH: 20284  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature



; LOCATION: (1)...(20284)  
; OTHER INFORMATION: n = a, t, c, or g  
US-09-526-193A-21

Alignment Scores: 447 Length: 20284  
Pred. No.: 9.00 Matches: 9  
Score: 100.00% Conservative: 0  
Percent Similarity: 100.00% Mismatches: 0  
Best Local Similarity: 100.00% Indels: 0  
Query Match: 0.95% Gaps: 0  
DB: 4

US-09-671-687A-3 (1-949) x US-09-526-193A-21 (1-20284)

Qy 521 LeuLysLysAlaLeuPheValLysLeu 529

Db 6166 CTTAAAGGCAATTATTGTGAACCTT 6192

#### RESULT 7

US-09-949-016-12572/c  
; Sequence 12572, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 12572  
; LENGTH: 53806  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-12572

Alignment Scores: 1.22e+03 Length: 53806  
Pred. No.: 9.00 Matches: 9  
Score: 100.00% Conservative: 0  
Percent Similarity: 100.00% Mismatches: 0  
Best Local Similarity: 100.00% Indels: 0  
Query Match: 0.95% Gaps: 0  
DB: 4

US-09-671-687A-3 (1-949) x US-09-949-016-12572 (1-53806)

Qy 691 ProLeuLysIleArgSerAlaGly 699

Db 42101 CCACTACTGAGATCGATCTCCGCGC 42075

#### RESULT 8

US-09-134-001C-2194/c  
; Sequence 2194, Application US/09134001C  
; Patent No. 6380370  
; GENERAL INFORMATION:  
; APPLICANT: Lynn Doucette-Stamm et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS  
; EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: GTC-007  
; CURRENT APPLICATION NUMBER: US/09/134,001C  
; CURRENT FILING DATE: 1998-08-13  
; PRIOR APPLICATION NUMBER: US 60/064,964  
; PRIOR FILING DATE: 1997-11-08  
; PRIOR APPLICATION NUMBER: US 60/055,779  
; PRIOR FILING DATE: 1997-08-14  
; NUMBER OF SEQ ID NOS: 5674

; SEQ ID NO 2194  
; LENGTH: 255  
; TYPE: DNA  
; ORGANISM: Staphylococcus epidermidis  
US-09-134-001C-2194

Alignment Scores: 57.3 Length: 255  
Pred. No.: 8.00 Matches: 8  
Score: 100.00% Conservative: 0  
Percent Similarity: 100.00% Mismatches: 0  
Best Local Similarity: 100.00% Indels: 0  
Query Match: 0.84% Gaps: 0  
DB: 3

US-09-671-687A-3 (1-949) x US-09-134-001C-2194 (1-255)

Qy 731 PheIleAsnSerAsnLeuLysPhe 738

Db 180 TTCATTACAGCAATTGTGAATTT 157

#### RESULT 9

US-09-232-785-240/c  
; Sequence 240, Application US/09232785  
; Patent No. 6733965  
; GENERAL INFORMATION:  
; APPLICANT: International Paper Co.  
; APPLICANT: Echt, Craig S  
; APPLICANT: Nelson, C. Dana  
; TITLE OF INVENTION: MICROSATELLITE DNA MARKERS AND USES  
; THEREOF  
; FILE REFERENCE: 4481/1E188US1  
; CURRENT APPLICATION NUMBER: US/09/232,785  
; CURRENT FILING DATE: 1999-01-19  
; PRIOR APPLICATION NUMBER: 09/232,884  
; PRIOR FILING DATE: 1999-01-15  
; NUMBER OF SEQ ID NOS: 397  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 240  
; LENGTH: 266  
; TYPE: DNA  
; ORGANISM: Pinus taeda L.  
US-09-232-785-240

Alignment Scores: 59.8 Length: 266  
Pred. No.: 8.00 Matches: 8  
Score: 100.00% Conservative: 0  
Percent Similarity: 100.00% Mismatches: 0  
Best Local Similarity: 100.00% Indels: 0  
Query Match: 0.84% Gaps: 0  
DB: 4

US-09-671-687A-3 (1-949) x US-09-232-785-240 (1-266)

Qy 649 AlaThrLysIleMetLysLeuArg 656

Db 38 GCGACAAAGATTATGAAGACTGAGA 15

#### RESULT 10

US-09-949-016-52259  
; Sequence 52259, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; NUMBER OF SEQ ID NOS: 60/231,498

; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 52259  
; LENGTH: 601  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-52259

Alignment Scores:  
Pred. No.: 138 Length: 601  
Score: 8.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 0.84% Indels: 0  
DB: 4 Gaps: 0

US-09-671-687A-3 (1-949) x US-09-949-016-52259 (1-601)

Qy 608 SerValleuAepThrValleuLeu 615  
Db 527 TCTGTTCTGATGCTGCTGCTA 550

## RESULT 11

US-09-949-016-52319/c  
; Sequence 52319, Application US/09949016  
; Patent No. 6812339

; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 52319

; LENGTH: 601  
; TYPE: DNA

; ORGANISM: Human  
US-09-949-016-52319

Alignment Scores:  
Pred. No.: 138 Length: 601  
Score: 8.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 0.84% Indels: 0  
DB: 4 Gaps: 0

US-09-671-687A-3 (1-949) x US-09-949-016-52319 (1-601)

Qy 2 SerSerGlyLeuTrpSerGlnGlu 9  
Db 554 TCCTCTGCTCTCTGCTCAGGAG 531

## RESULT 12

US-09-949-016-52320/c  
; Sequence 52320, Application US/09949016  
; Patent No. 6812339

; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 52319

; LENGTH: 601  
; TYPE: DNA

; ORGANISM: Human  
US-09-949-016-52319

; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 52320  
; LENGTH: 601  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-52320

Alignment Scores:  
Pred. No.: 138 Length: 601  
Score: 8.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 0.84% Indels: 0  
DB: 4 Gaps: 0

US-09-671-687A-3 (1-949) x US-09-949-016-52320 (1-601)

Qy 2 SerSerGlyLeuTrpSerGlnGlu 9  
Db 525 TCCTCTGCTCTCTGCTCAGGAG 502

## RESULT 13

US-09-949-016-52321/c  
; Sequence 52321, Application US/09949016  
; Patent No. 6812339

; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307

; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 52321

; LENGTH: 601  
; TYPE: DNA

; ORGANISM: Human  
US-09-949-016-52321

Alignment Scores:  
Pred. No.: 138 Length: 601  
Score: 8.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 0.84% Indels: 0  
DB: 4 Gaps: 0

US-09-671-687A-3 (1-949) x US-09-949-016-52321 (1-601)

Qy 2 SerSerGlyLeuTrpSerGlnGlu 9  
Db 158 TCCTCTGCTCTCTGCTCAGGAG 135

## RESULT 14

US-09-949-016-56429/c  
; Sequence 56429, Application US/09949016  
; Patent No. 6812339

; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.

Search completed: April 17, 2005, 11:43:59  
Job time : 429 secs

US-09-671-687A-3 (1-949) X US-09-949-016-56429 (1-601)

RESULT 15

```

/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
/ WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
/
/ FILE OF INVENTION: CU001307
/
/ CURRENT APPLICATION NUMBER: US/09/949,016
/
/ CURRENT FILING DATE: 2000-04-14
/
/ PRIOR APPLICATION NUMBER: 60/241,755
/
/ PRIOR FILING DATE: 2000-10-20
/
/ PRIOR APPLICATION NUMBER: 60/237,768
/
/ PRIOR FILING DATE: 2000-10-03
/
/ PRIOR APPLICATION NUMBER: 60/231,498
/
/ PRIOR FILING DATE: 2000-09-08
/
/ NUMBER OF SEQ ID NOS: 207012
/
/ SOFTWARE: FastSeq for Windows Version 4.0
/
/ SEQ ID NO 56450
/

```

US-09-671-687A-3 (1-949) x US-09-949-016-56450 (1-601)

<b>Qy</b>	391 SerSerProLeuGlnProPro	398
<b>Db</b>	579 TCCTCTCCTCCCCTCCAACTCCG	556

```

-----
CC EMBL; D26185; BAA05202.1; -
DR EMBL; Z99124; CAB16108.1; -
DR PIR; S65996; S65996.
DR Subtilist; BG10030; Yyba.
DR InterPro; IPR002197; HTH_Fis.
DR InterPro; IPR000835; HTH_MarR.
DR InterPro; IPR009058; Wing_hlx_DNA_bnd.
DR Pfam; PF01047; MarR; 1.
DR PRINTS; PR00598; HTHMARR.
DR SMART; SMO0347; HTH_MARR; 1.
DR TIGRFAMs; TIGR01199; HTH_fis; 1.
DR PROSITE; PS01117; HTH_MARR_1; 1.
DR PROSITE; PS00995; HTH_MARR_2; 1.
KW Complete proteome; DNA-binding; Hypothetical protein;
KW Transcription regulation.
FT DOMAIN 1 133 HTH marR-type.
SQ SEQUENCE 150 AA; 17584 MW; E9195BBAD02DF4DC CRC64;

Query Match 0.8%; Score 8; DB 1; Length 150;
Best Local Similarity 100.0%; Pred.No. 37;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 666 SGTSEEK 673
Db 114 SGTSEEK 121

RESULT 15
Q7QYI6
ID Q7QYI6 PRELIMINARY; PRT; 212 AA.
AC Q7QYI6;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE GLP 80 31039 31677.
OS Giardia lamblia ATCC 50803.
OC Eukaryota; Diplomonadida; Hexamitidae; Giardiinae; Giardia.
OX NCBI_TaxID=184922;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=WB C6;
RA Morrison H.G., McArthur A.G., Adam R.D., Aley S.B., Gillin F.D.,
RA Olsen G.J., Sogin M.L.;
RT "Draft sequence of the Giardia lamblia genome.";
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AACB0100050; EAA40124.1; -
SQ SEQUENCE 212 AA; 24313 MW; ECE7494F74925675 CRC64;

Query Match 0.8%; Score 8; DB 2; Length 212;
Best Local Similarity 100.0%; Pred.No. 51;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 196 ALDKLELI 203
Db 103 ALDKLELI 110

```

Search completed: April 18, 2005, 09:40:35  
Job time : 70 secs

```
RESULT 12
Q6AQ83          PRELIMINARY;      PRT;   225 AA.
ID  Q6AQ83;
AC  Q6AQ83;
DT  25-OCT-2004 (TRENBLrel. 28, Created)
DT  25-OCT-2004 (TRENBLrel. 28, Last sequence update)
DT  25-OCT-2004 (TRENBLrel. 28, Last annotation update)
DE  Related to xylene monooxygenase electron transfer component.
GN  OrderedLocusNames=DP0761;
OS  Desulfotalea psychrophila.
OC  Bacteria; Proteobacteria; Deltaproteobacteria; Desulfobacteriales;
OC  Desulfobulbaceae; Desulfotalea.
OX  NCBI_TaxID=84980;
[1]
SEQUENCE FROM N.A.
RP  STRAIN=LSV54 / DSM 12343;
RC  PubMed=15305914;
RX  Rabus R., Ruepp A., Frickey T., Rattei T., Fartmann B., Stark M.,
RA  Bauer M., Zibat A., Lombardot T., Becker I., Amann J., Gellner K.,
RA  Teeling H., Leuschner W.D., Gloeckner F.-O., Lupas A.N., Amann R.,
RA  Klenk H.-P.;
RT  "The genome of Desulfotalea psychrophila, a sulfate-reducing bacterium
RT  from permanently cold Arctic sediments."
RL  Environ. Microbiol. 6:897-902(2004).
DR  EMBL; CR522870; CAG35490.1; -.
DR  GO; GO:0016491; F:oxidoreductase activity; IEA.
DR  GO; GO:0006118; P:electron transport; IEA.
DR  InterPro; IPR001834; Cyt_B5_reductase.
DR  InterPro; IPR001709; FPN_cyt_reductase.
DR  InterPro; IPR001433; Oxred_FAD/NAD(P).
DR  Pfam; PF00175; NAD binding 1; 1.
DR  PRINTS; PR00406; CYTB5RDTASE.
DR  PRINTS; PR00371; FPNCR.
DR  PRINTS; PR00410; PHEHYDRLASE.
DR  Complete proteome.
SQ  SEQUENCE 225 AA; 24909 MW; D5B2FA62F7A3FE4 CRC64;

Query Match          0.9%; Score 9; DB 2; Length 225;
Best Local Similarity 100.0%; Pred. No. 4.7;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy  41 VPKGSIGQY 49
Db  183 VPKGSIGQY 191

RESULT 13
Q8EX39          PRELIMINARY;      PRT;   80 AA.
ID  Q8EX39;
AC  Q8EX39;
DT  01-MAR-2003 (TRENBLrel. 23, Created)
DT  01-MAR-2003 (TRENBLrel. 23, Last sequence update)
DT  01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE  Hypothetical protein MYPE110.
GN  OrderedLocusNames=MYPE110;
OS  Mycoplasma penetrans.
OC  Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX  NCBI_TaxID=28227;
[1]
SEQUENCE FROM N.A.
RP  STRAIN=HF-2;
RC  MEDLINE=22354719; PubMed=12466555; DOI=10.1093/nar/gkf667;
RX  Sasaki Y., Ishikawa J., Yamashita A., Oshima K., Kenri T.,
RA  Yoshino C., Horino A., Shiba T., Sasaki T., Hattori M.;
RT  "The complete genomic sequence of Mycoplasma penetrans, an
RT  intracellular bacterial pathogen in humans."
RL  Nucleic Acids Res. 30:5293-5300(2002).
DR  EMBL; AP004170; BAC43801.1; -.
KW  Complete proteome; Hypothetical protein.
SQ  SEQUENCE 80 AA; 9231 MW; 42888D13B01030DA CRC64;

Query Match          0.8%; Score 8; DB 2; Length 80;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy  197 LDKLELIE 204
Db  32 LDKLELIE 39

RESULT 14
YVBA_BACSU
ID  YVBA_BACSU STANDARD;      PRT;   150 AA.
AC  P37503;
DT  01-OCT-1994 (Rel. 30, Created)
DT  01-OCT-1994 (Rel. 30, Last sequence update)
DT  25-OCT-2004 (Rel. 45, Last annotation update)
DE  Putative HTH-type transcriptional regulator yvba.
GN  Name=yvba; OrderedLocusNames=BSU40710;
OS  Bacillus subtilis.
OC  Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX  NCBI_TaxID=1423;
[1]
SEQUENCE FROM N.A.
RP  STRAIN=168;
RC  MEDLINE=96051385; PubMed=7584024;
RX  Ogasawara N., Nakai S., Yoshikawa H.;
RT  "Systematic sequencing of the 180 kilobase region of the Bacillus
RT  subtilis chromosome containing the replication origin."
RL  DNA Res. 1:1-14(1994).
[2]
SEQUENCE FROM N.A.
RP  STRAIN=168;
RC  MEDLINE=98044033; PubMed=9384377; DOI=10.1038/36786;
RX  Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA  Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA  Borriess R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA  Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA  Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
RA  Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,
RA  Entian K.-D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA  Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA  Ghim S.Y., Glaser P., Goffeau A., Gollightly E.J., Grand G.,
RA  Guisepi G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
RA  Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M.,
RA  Jones L.-M., Joris B., Karamata D., Kasahara Y., Klaert-Blanchard M.,
RA  Klein C., Kobayashi Y., Koetter P., Koningsstein G., Krogh S.,
RA  Kumano M., Kurita K., Lapidus A., Lardinois S., Lauber J.,
RA  Lazarevic V., Lee S.M., Levine A., Liu H., Masuda S., Mauel C.,
RA  Medigue C., Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S.,
RA  Noback M., Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B.,
RA  Park S.H., Parro V., Pohl T.M., Portetelle D., Porwollik S.,
RA  Prescott A.M., Prescan E., Pujic P., Purnelle B., Rapoport G.,
RA  Rey M., Reynolds S., Rieger M., Rivolta C., Rocha E., Roche B.,
RA  Rose M., Sadaie Y., Sato T., Scanlan E., Schleich S., Schroeter R.,
RA  Scoffone F., Sekiguchi J., Sekowska A., Seror S.J., Seror P.,
RA  Shin B.S., Soldo B., Sorokin A., Tacconi E., Takagi T., Takahashi H.,
RA  Takemaru K., Takeuchi M., Tanakoshi A., Tanaka T., Terpstra P.,
RA  Togononi A., Tosato V., Uchiyama S., Vandenbol M., Vannier F.,
RA  Vassaretti A., Viari A., Wambutt R., Wedler E., Wedler H.,
RA  Weitzensogger T., Winters P., Wipat A., Yamamoto H., Yamane K.,
RA  Yasumoto K., Yata K., Yoshida K., Yoshikawa H.F., Zumstein E.,
RA  Yoshikawa H., Danchin A.;
RT  "The complete genome sequence of the Gram-positive bacterium Bacillus
RT  subtilis."
RL  Nature 390:249-256(1997).
CC  -!- SIMILARITY: Contains 1 HTH marR-type DNA-binding domain.
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
```

OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;  
 OC Rhabditidae; Peloderae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RX MEDLINE=99069613; PubMed=9851916;  
 none;  
 RA "Genome sequence of the nematode C.elegans: A platform for  
 RT investigating biology.";  
 RL Science 282:2012-2018(1998).  
 [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RA Kershaw J.K.;  
 RL Submitted (FEB-1994) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; Z30215; CAF31477.2; -;  
 DR WormBase; WBGene0009594; F40F12.5.  
 DR WormPep; F40F12.5; CE36647.  
 DR InterPro; IPR000938; CAP-Gly.  
 DR Pfam; PF01302; CAP\_GLY; 1.  
 KW Hypothetical protein.  
 SQ SEQUENCE 1144 AA; 128194 MW; 33CCDB2ED3997FE7 CRC64;  
 Query Match 1.3%; Score 12; DB 2; Length 1144;  
 Best Local Similarity 100.0%; Pred.No. 0.014;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 856 AVLCIETSHYVA 867  
 Db 1026 AVLCIETSHYVA 1037  
 RESULT 11  
 Q9VL04 PRELIMINARY; PRT; 517 AA.  
 AC Q9VL04;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-MAY-2004 (TrEMBLrel. 26, Last annotation update)  
 DE CG5603-PA.  
 GN ORFNames=CG5603;  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;  
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Anantides P.G., Scher S.E., Li P.W., Hoskins R.A., Galie R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.H., Blazei R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.B.,  
 RA Abril J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,  
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Dou L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 RA Fosler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodek A., Gong P., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,  
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,

RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,  
 RA Shie B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.Y., Wassarman D.A., Weinstein G.M., Weisenbach J.,  
 RA Williams S.M., Woodagef, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,  
 RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of Drosophila melanogaster.";  
 RL Science 287:2185-2195(2000).  
 [2]  
 RN SEQUENCE FROM N.A.  
 RX MEDLINE=22426065; PubMed=12537568;  
 RA Celniker S.E., Wheeler D.A., Krommiller B., Carlson J.W., Halpern A.,  
 RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,  
 RA George R.A., Hoskins R.A., Laverty T., Muzny D.M., Nelson C.R.,  
 RA Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,  
 RA Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,  
 RA Weinstock G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;  
 RT "Finishing a whole-genome shotgun: Release 3 of the Drosophila  
 RT melanogaster euchromatic genome sequence";  
 RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).  
 [3]  
 RN SEQUENCE FROM N.A.  
 RX MEDLINE=22426070; PubMed=12537573;  
 RA Kaminker J.S., Bergman C.M., Krommiller B., Carlson J., Svirskas R.,  
 RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,  
 RA Ashburner M., Celniker S.E.;  
 RT "The transposable elements of the Drosophila melanogaster euchromatin:  
 RT a genomics perspective";  
 RL Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).  
 [4]  
 RN SEQUENCE FROM N.A.  
 RX MEDLINE=22426069; PubMed=12537572;  
 RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,  
 RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,  
 RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,  
 RA Bettencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,  
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,  
 RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
 RA Lewis S.E.;  
 RT "Annotation of the Drosophila melanogaster euchromatic genome: a  
 RT systematic review";  
 RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).  
 [5]  
 RN SEQUENCE FROM N.A.  
 RG FlyBase;  
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.  
 [6]  
 RN SEQUENCE FROM N.A.  
 RG FlyBase;  
 RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AE003628; AAF52901.1; -;  
 DR FlyBase; FBgn0032210; CG5603.  
 DR InterPro; IPR000938; CAP-Gly.  
 DR Pfam; PF01302; CAP\_GLY; 1.  
 SQ SEQUENCE 517 AA; 58081 MW; AESE9802F2989404 CRC64;  
 Query Match 1.1%; Score 10; DB 2; Length 517;  
 Best Local Similarity 100.0%; Pred.No. 0.88;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 671 EKDPEEFILN 680  
 Db 361 EKDPEEFILN 370



RA [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;  
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
RA Brandon R.C., Rogers Y.H., Blazej R.G., Champe M., Pfeiffer B.D.,  
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,  
RA Abril J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,  
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
RA Borkova D., Botchan M.R., Bouck J., Brockstein P., Brotier P.,  
RA Burtis K.C., Busam D.A., Butler H., Cadiou E., Center A., Chandra I.,  
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,  
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Lasako P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,  
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reaenert K., Remington K., Saunders R.D., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Kianos I., Simpson M., Skupski M.P., Smith T.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.Y., Wasarman D.A., Weinstein G.M., Weissenbach J.,  
RA Williams S.M., Woodagef, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,  
RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RT "The genome sequence of *Drosophila melanogaster*";  
RL Science 287:2185-2195(2000).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=22426065; PubMed=12537568;  
RA Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,  
RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,  
RA George R.A., Hoskins R.A., Laverty T., Muzny D.M., Nelson C.R.,  
RA Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,  
RA Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,  
RA Weinstein G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;  
RT "Finishing a whole-genome shotgun: Release 3 of the *Drosophila*  
RT *melanogaster* euchromatic genome sequence";  
RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).  
RN [3]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=22426070; PubMed=12537573;  
RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J., Svirskas R.,  
RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,  
RA Ashburner M., Celniker S.E.;  
RT "The transposable elements of the *Drosophila melanogaster* euchromatin:  
RT a genomics perspective";  
RL Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).  
RN [4]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=22426069; PubMed=12537572;  
RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,  
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,  
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,  
RA Bettencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,  
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.O.,  
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
RA Lewis S.E.;  
RT "Annotation of the *Drosophila melanogaster* euchromatic genome: a

RT systematic review.";  
RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).  
RN [5]  
RP SEQUENCE FROM N.A.  
RG FlyBase;  
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.  
RN [6]  
RP SEQUENCE FROM N.A.  
RG FlyBase;  
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AS003628; AAN10738.1; -  
DR FlyBase; FBgn0032210; CG5603.  
DR InterPro; IPR000938; CAP-Gly.  
DR Pfam; PF01302; CAP\_Gly; 1.  
SQ SEQUENCE 639 AA; 72128 MW; 6E29D3B09FF5E55B CRC64;  
Query Match 1.3%; Score 12; DB 2; Length 639;  
Best Local Similarity 100.0%; Pred. No. 0.0083;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 859 CIETSHYVAFVK 870  
DB 545 CIETSHYVAFVK 556  
|||||  
RESULT 9  
Q8SYF0 PRELIMINARY; PRT; 639 AA.  
AC Q8SYF0;  
DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
DT 01-JAN-2004 (TrEMBLrel. 26, Last annotation update)  
DE RE64280P.  
GN ORFNames=CG5603;  
OS *Drosophila melanogaster* (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN [1]\_TaxID=7227;  
RP SEQUENCE FROM N.A.  
RC STRAIN=Berkley;  
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,  
RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,  
RA George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,  
RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Faragas V., Park S.,  
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,  
RA Celniker S.;  
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY071592; AAL49214.1; -  
DR FlyBase; FBgn0032210; CG5603.  
DR InterPro; IPR000938; CAP-Gly.  
DR Pfam; PF01302; CAP\_Gly; 1.  
SQ SEQUENCE 639 AA; 72162 MW; 6E29D3B8B757675B CRC64;  
Query Match 1.3%; Score 12; DB 2; Length 639;  
Best Local Similarity 100.0%; Pred. No. 0.0083;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 859 CIETSHYVAFVK 870  
DB 545 CIETSHYVAFVK 556  
|||||  
RESULT 10  
Q7JMS4 PRELIMINARY; PRT; 1144 AA.  
AC Q7JMS4;  
DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
DE Hypothetical protein F40F12.5.  
GN ORFNames=F40F12.5;

```

RN  [6]
RP  SEQUENCE FROM N.A.
RG  FlyBase;
RL  Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
DR  EMBL; AE003628; AAN10740.1; -.
DR  FlyBase; FBgn0032210; CG5603.
DR  InterPro; IPR000938; CAP-Gly.
DR  Pfam; PF01302; CAP_GLY; 1.
SQ  SEQUENCE 550 AA; 62083 MW; C87C73651D88FD59 CRC64;

Query Match      1.3%; Score 12; DB 2; Length 550;
Best Local Similarity 100.0%; Pred. No. 0.0072;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  859 CIETSHYVAFVK 870
Db  456 CIETSHYVAFVK 467
      |||||
      |||||

RESULT 7
Q81PC3 PRELIMINARY; PRT; 551 AA.
AC  Q81PC3;
DT  01-MAR-2003 (TrEMBLrel. 23, Created)
DT  01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT  01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE  CG5603-PB.
GN  ORFNames=CG5603;
OS  Drosophila melanogaster (Fruit fly).
OC  Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC  Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC  Ephydroidea; Drosophilidae; Drosophila.
OX  NCBI_TaxID=7227;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA  Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA  Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA  Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA  Brandon R.C., Rogers Y.H., Blazej R.G., Champe M., Pfeiffer B.D.,
RA  Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
RA  Abrial J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
RA  Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA  Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA  Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA  Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA  Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA  de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA  Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA  Durbin K.J., Evangelista C.C., Ferraz C., Ferrier S., Fleischmann W.,
RA  Foster K.C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA  Glisdek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA  Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA  Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,
RA  Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA  Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA  Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA  Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA  Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA  Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA  Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacle J.M.,
RA  Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA  Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
RA  Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA  Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA  Svirskaas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA  Wang Z.Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA  Williams S.M., Woodger, Morley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
RA  Yen R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA  Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Smith H.O.,
RA  Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT  "The genome sequence of Drosophila melanogaster.";
```

```

RL  Science 287:2185-2195(2000).
RN  [2]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=22426065; PubMed=12537568;
RA  Celniker S.E., Wheeler D.A., Krommiller B., Carlson J.W., Halpern A.,
RA  Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
RA  George R.A., Hoskins R.A., Laverty T., Muzny D.M., Nelson C.R.,
RA  Pacle J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
RA  Svirskaas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
RA  Weinstock G., Scher S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
RT  "Finishing a whole-genome shotgun: Release 3 of the Drosophila
RT  melanogaster euchromatic genome sequence.";
```

Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).

```

RN  [3]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=22426070; PubMed=12537573;
RA  Kaminker J.S., Bergman C.M., Krommiller B., Carlson J., Svirskaas R.,
RA  Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
RA  Ashburner M., Celniker S.E.;
RT  "The transposable elements of the Drosophila melanogaster euchromatin:
RT  a genomics perspective.";
```

Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).

```

RN  [4]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=22426069; PubMed=12537572;
RA  Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA  Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
RA  Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA  Bettencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,
RA  Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA  Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA  Lewis S.E.;
RT  "Annotation of the Drosophila melanogaster euchromatic genome: a
RT  systematic review.";
```

Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).

```

RN  [5]
RP  SEQUENCE FROM N.A.
RG  FlyBase;
RL  Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
RN  [6]
RP  SEQUENCE FROM N.A.
RG  FlyBase;
RL  Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
DR  EMBL; AE003628; AAN10741.1; -.
DR  MEROPS; C67.001; -.
DR  FlyBase; FBgn0032210; CG5603.
DR  InterPro; IPR000938; CAP-Gly.
DR  Pfam; PF01302; CAP_GLY; 1.
SQ  SEQUENCE 551 AA; 62315 MW; 9D4D63B2287B7783 CRC64;

Query Match      1.3%; Score 12; DB 2; Length 551;
Best Local Similarity 100.0%; Pred. No. 0.0072;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  859 CIETSHYVAFVK 870
Db  457 CIETSHYVAFVK 468
      |||||
      |||||

RESULT 8
Q81PC5 PRELIMINARY; PRT; 639 AA.
AC  Q81PC5;
DT  01-MAR-2003 (TrEMBLrel. 23, Created)
DT  01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT  01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE  CG5603-PD (CG5603-pe).
GN  ORFNames=CG5603;
OS  Drosophila melanogaster (Fruit fly).
OC  Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC  Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC  Ephydroidea; Drosophilidae; Drosophila.
OX  NCBI_TaxID=7227;
```

```

Matches 98; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 852 MELFAVLCIETSHYVAFVKYKDDSAWLFDFSDMADRGQGNFIPQVTFPCPEVGYLKM 911
Db 357 MELFAVLCIETSHYVAFVKYKDDSAWLFDFSDMADRGQGNFIPQVTFPCPEVGYLKM 416
QY 912 SLEDLHSLDSRRIOGCARLLCDAYCMYQSTMSLYK 949
Db 417 SLEDLHSLDSRRIOGCARLLCDAYCMYQSTMSLYK 454

RESULT 5
Q7Q414 PRELIMINARY; PRT; 778 AA.
AC Q7Q414;
DT 01-MAR-2004 (TREMBLrel. 26, Created)
DT 01-MAR-2004 (TREMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE AGCP11527 (fragment).
GN Name=agCG56065; ORFNames=ENSGG00000016806;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Anopheles.
OX NCBI_TaxID=180454;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PEST;
RA Anopheles Genome Sequencing Consortium;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAB01008964; EAA12869.1; -.
DR MEROPS; C67.001; -.
DR GO; GO:0004197; F:cytosteine-type endopeptidase activity; IEA.
DR GO; GO:0004221; F:ubiquitin thiolesterase activity; IEA.
DR GO; GO:0006511; P:ubiquitin-dependent protein catabolism; IEA.
DR InterPro; IPR000938; CAP-Gly.
DR Pfam; PF01302; CAP_Gly; 1.
DR PROSITE; PS02335; UCH_2_3; 1.
FT NON TER 1
SQ SEQUENCE 778 AA; 83981 MW; 0A775088B20864A7 CRC64;

Query Match 1.4%; Score 13; DB 2; Length 778;
Best Local Similarity 100.0%; Fred. No. 0.00087;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 742 PSCLIIQMPRFK 754
Db 568 PSCLIIQMPRFK 580

RESULT 6
Q8IPC4 PRELIMINARY; PRT; 550 AA.
AC Q8IPC4;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE CG5603-PC.
OS ORFNames=CG5603;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20196006; PubMed=107311132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,

```

```

RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.H., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
RA Abril J.F., Aghayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballwey R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrier S., Fleischmann W.,
RA Foslter C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclab J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodgett, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RA "The genome sequence of Drosophila melanogaster.";
Science 287:2185-2195 (2000).
RL [2]
RN SEQUENCE FROM N.A.
RX MEDLINE=22426065; PubMed=12537568;
RA Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,
RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
RA George R.A., Hoskins R.A., Laverly T., Muzny D.M., Nelson C.R.,
RA Paclab J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
RA Svirskas R., Taber P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
RA Weinstein G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
RT "Finishing a whole-genome shotgun: Release 3 of the Drosophila
RT melanogaster euchromatic genome sequence.";
Genome Biol. 3:RESEARCH0079-RESEARCH0079 (2002).
RL [3]
RN SEQUENCE FROM N.A.
RX MEDLINE=22426070; PubMed=12537573;
RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J., Svirskas R.,
RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
RA Ashburner M., Celniker S.E.;
RT "The transposable elements of the Drosophila melanogaster euchromatin:
RT a Genomics perspective.";
Genome Biol. 3:RESEARCH0084-RESEARCH0084 (2002).
RL [4]
RN SEQUENCE FROM N.A.
RX MEDLINE=22426069; PubMed=12537572;
RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA Bettencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RT "Annotation of the Drosophila melanogaster euchromatic genome: a
RT systematic review.";
Genome Biol. 3:RESEARCH0083-RESEARCH0083 (2002).
RL [5]
RN SEQUENCE FROM N.A.
RG FlyBase;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.

```



```
Db 471 EVGLAEYKRPFPFVGIRWIGQPPGLNEVLAGELEDEACAGTGTGRGTGYFTCALKK 530
Qy 524 ALFVKLKCRDPSRFPASQVNSQIERCNSLAFGGYLSVEENTPPKMEKEGLEIMIGK 583
Db 531 ALFVKLKCRDPSRFPASQVNSQIERCNSLAFGGYLSVEENTPPKMEKEGLEIMIGK 590
Qy 584 KKGIOGHVNSCYLDSTLFCCLFAFSSVLDVTLRLPKRKNDVEYYSETQELLRTIENVPLRI 643
Db 591 KKGIOGHVNSCYLDSTLFCCLFAFSSVLDVTLRLPKRKNDVEYYSETQELLRTIENVPLRI 650
Qy 644 YGYVCATKIMKRLKLEKVEAASGTSSEKDEEFLNLFHILRVEPLLKIRSAGQKQV 703
Db 651 YGYVCATKIMKRLKLEKVEAASGTSSEKDEEFLNLFHILRVEPLLKIRSAGQKQV 710
Qy 704 DCYFQIPEWKEKVGVPITQQLLWSFNSLNKFAEAPSCLI IOMPRGKDFKLFKTF 763
Db 711 DCYFQIPEWKEKVGVPITQQLLWSFNSLNKFAEAPSCLI IOMPRGKDFKLFKTF 770
Qy 764 PSLELNITDLEDTPQRCICGLAMYECCYDPPDISAGKIKQFCCKTQNTQVHLHPKR 823
Db 771 PSLELNITDLEDTPQRCICGLAMYECCYDPPDISAGKIKQFCCKTQNTQVHLHPKR 830
Qy 824 LNHNKPNVSLPKDLPDWRHGCIPQNNMELFAVLICFETSHYVAFVKYKDDSAWLFDD 883
Db 831 LNHNKPNVSLPKDLPDWRHGCIPQNNMELFAVLICFETSHYVAFVKYKDDSAWLFDD 890
Qy 884 MADRGGGNGFNIPQVTPCPPEVGEVYKMSLEDLHSLDSRRIOGCARRLLCDAYMCYQSP 943
Db 891 MADRGGGNGFNIPQVTPCPPEVGEVYKMSLEDLHSLDSRRIOGCARRLLCDAYMCYQSP 950
Qy 944 TMSLYK 949
Db 951 TMSLYK 956

RESULT 2
CYLD_MOUSE STANDARD; PRT: 952 AA.
AC Q80TQ2; Q80VB3; Q8BX23; Q8BYL9; Q8CGB0;
DT 25-OCT-2004 (Rel. 45, Created)
DT 25-OCT-2004 (Rel. 45, Last sequence update)
DE Probable ubiquitin carboxyl-terminal hydrolase CYLD (EC 3.1.2.15)
DE (Ubiquitin thiolesterase CYLD) (Ubiquitin-specific processing protease
DE CYLD) (Deubiquitinating enzyme .CYLD).
GN Name=Cyld; Synonyms=Cyld1, Kiaa0849;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RC TISSUE=Brain;
RX MEDLINE=22579291; PubMed=12693553;
RA Okazaki N., Kikuno R., Inamoto S., Aizawa H., Yuasa S.,
RA Nakajima D., Nagase T., Ohara O., Koga H.;
RT "Prediction of the coding sequences of mouse homologues of KIAA gene:
RT II. The complete nucleotide sequences of 400 mouse KIAA-homologous
RT cDNAs identified by screening of terminal sequences of cDNA clones
RT randomly sampled from size-fractionated libraries.";
RL DNA Res. 10:35-48(2003).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 3), AND SEQUENCE OF 1-620 FROM N.A.
RP (ISOFORM 1).
RC STRAIN=C57BL/6J; TISSUE=Cerebellum, and Hypothalamus;
RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA Nikaïdo I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
```

```
RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA Grimmond S., Gustinchik S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA Ravasi T., Reed J.C., Reid D.J., Reid J., Ring B.Z., Ringwald M.,
RA Sandelin A., Schneider C., Sempile C.A., Setou M., Shimada K.,
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
RA Wilming L.G., Wyshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA Hara A., Hashizume W., Imotani K., Iehii Y., Itoh M., Kagawa I.,
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinegawa A.,
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA Birney E., Hayashizaki Y.;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573 (2002).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RC STRAIN=FVB/N; TISSUE=Mammary gland;
RX MEDLINE=2388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max A.A., Wang J., Hong F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smails D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
CC -!- FUNCTION: Negative regulator of TRAF2 and NF-kappa-B signaling
CC pathway. Has deubiquitinating activity that is directed towards
CC non-Lys-48-linked polyubiquitin chains. The inhibition of NF-
CC kappa-B activation is mediated at least in part, by the
CC deubiquitination and inactivation of TRAF2 and, to a lesser
CC extent, TRAF6 (By similarity).
CC -!- CATALYTIC ACTIVITY: Ubiquitin C-terminal thiolester + H2O =
CC ubiquitin + a thiol.
CC -!- SUBUNIT: Interacts with NEMO, TRAF2 and TRIP (By similarity).
CC -!- SUBCELLULAR LOCATION: Perinuclear region (By similarity).
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=3;
CC Name=2;
CC IsoID=Q80TQ2-1; Sequence=Displayed;
CC Name=1;
CC IsoID=Q80TQ2-2; Sequence=VSP_011278;
CC Name=3;
CC IsoID=Q80TQ2-3; Sequence=VSP_011279, VSP_011280;
CC -!- SIMILARITY: Belongs to the peptidase C67 family.
CC -!- SIMILARITY: Contains 2 CAP-Gly domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
```





GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 18, 2005, 09:15:43 ; Search time 65 Seconds  
(without alignments)  
7476.356 Million cell updates/sec

Title: US-09-671-687A-3  
Perfect score: 949  
Sequence: 1 MSSGLMSQEKVTSYPWEERI.....RLLCDAYMCMYQSPMTSLYK 949

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 1612378 seqs, 512079187 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : UniProt\_03: \*  
1: uniprot\_sprot: \*  
2: uniprot\_trembl: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	606	63.9	956	1	CYLD_HUMAN
2	116	12.2	952	1	CYLD_MOUSE
3	116	12.2	953	2	Q6GH62
4	98	10.3	454	2	Q6TXJ6
5	13	1.4	778	2	Q7Q414
6	12	1.3	550	2	Q81PC4
7	12	1.3	551	2	Q81PC3
8	12	1.3	639	2	Q81PC5
9	12	1.3	639	2	Q8SYF0
10	12	1.3	1144	2	Q7JMS4
11	10	1.1	517	2	Q9VL04
12	9	0.9	225	2	Q6AQ83
13	8	0.8	80	2	Q8EX39
14	8	0.8	150	1	YVBA_BACSU
15	8	0.8	212	2	Q7QYI6
16	8	0.8	215	2	Q9HX43
17	8	0.8	216	2	Q9GKF5
18	8	0.8	225	2	Q7SFV2
19	8	0.8	236	2	Q9U0S3
20	8	0.8	237	2	Q8VKG1
21	8	0.8	258	2	Q7FJA3
22	8	0.8	274	2	Q62LQ0
23	8	0.8	276	2	Q63WB0
24	8	0.8	353	2	Q9EZU8
25	8	0.8	354	2	Q8N4A4
26	8	0.8	366	2	Q6A178
27	8	0.8	371	2	Q9AWT2
28	8	0.8	401	2	Q6MJH4
29	8	0.8	418	2	Q9VHV0
30	8	0.8	423	2	O01754
31	8	0.8	450	1	G6PI_LISIN

32	8	0.8	450	1	G6PI_LISMO
33	8	0.8	450	2	Q7IX61
34	8	0.8	458	2	O44617
35	8	0.8	522	2	Q6NAC1
36	8	0.8	524	1	CP11_MESAU
37	8	0.8	551	1	AG12_YEAST
38	8	0.8	555	2	Q818V0
39	8	0.8	608	1	ACEK_RALSO
40	8	0.8	625	2	Q8K7W4
41	8	0.8	651	2	Q912B2
42	8	0.8	740	2	O28905
43	8	0.8	753	2	Q8F356
44	8	0.8	760	2	Q9RXG0
45	8	0.8	763	2	Q72SG2

ALIGNMENTS

RESULT 1  
CYLD\_HUMAN  
ID CYLD\_HUMAN STANDARD: PRT: 956 AA.  
AC Q9NCQ7; Q94934; Q7L3N6; Q96EH0; Q9NZX9;  
DT 25-OCT-2004 (Rel. 45, Created)  
DT 25-OCT-2004 (Rel. 45, Last sequence update)  
DT 25-OCT-2004 (Rel. 45, Last annotation update)  
DE Probable ubiquitin carboxyl-terminal hydrolase CYLD (EC 3.1.2.15)  
DE (Ubiquitin thiolesterase CYLD) (Ubiquitin-specific processing protease  
DE CYLD) (Deubiquitinating enzyme CYLD) (HSPC057).  
GN Name=CYLD; Synonyms=CYLD1, KIAA0849;  
OS Homo sapiens (Human)  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A. (ISOFORM 1), DISEASE, AND TISSUE SPECIFICITY.  
RX MEDLINE=20296617; PubMed=10835629; DOI=10.1038/76006;  
RA Bignell G.R., Brown C., Biggs P.J., Lakhani S.R., Jones C., Hansen J.,  
RA Blair E., Hofmann B., Siebert R., Turner G., Evans D.G.,  
RA Schrandt-Stumpel C., Beemer F.A., Van Den Ouweland A., Halley D.,  
RA Dellpech B., Cleveland M.G., Leigh I., Leisti J., Rasmussen S.,  
RA Wallace M.R., Fenske C., Banerjee P., Oiso N., Chaggar R., Merritt S.,  
RA Leonard N., Huber M., Hohl D., Chapman P., Burn J., Swift S.,  
RA Smith A., Ashworth A., Stratton M.R.;  
RT "Identification of the familial cylindromatosis tumor suppressor  
RT gene.";  
RL Nat. Genet. 25:160-165(2000).  
RN [2]  
RP SEQUENCE FROM N.A. (ISOFORM 2).  
RX MEDLINE=99156230; PubMed=10048485;  
RA Nagase T., Ishikawa K.-I., Suyama M., Kikuno R., Hirose M.,  
RA Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;  
RT "Prediction of the coding sequences of unidentified human genes. XII.  
RT The complete sequences of 100 new cDNA clones from brain which code  
RT for large proteins in vitro.";  
RL DNA Res. 5:355-364(1998).  
RN [3]  
RP SEQUENCE FROM N.A. (ISOFORM 2).  
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altshuler S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
FAhey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,

QY	464	EVGSLAEVKENPPFVGVIIRWIGQPGLENEVLAGLELEDEACGCTDGTFRGTRYFTCALKK	523
Db	445	EVGSLAEVKENPPFVGVIIRWIGQPGLENEVLAGLELEDEACGCTDGTFRGTRYFTCALKK	504
QY	524	ALFVKLSCRDPDSRFASIQPVSNQIERCNSLAFGGYLSVWVENTPP	570
Db	505	ALFVKLSCRDPDSRFASIQPVSNQIERCNSLAFGGYLSVWVENTPP	551

RESULT 15

ABB89233 ID ABB89233 standard; protein: 261 AA.

AC ABB89233;

24-MAY-2002 (first entry)

XX  
DE Human polypeptide SEQ ID NO 1609.

Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;  
KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antitumor;  
KW vulnary; anticonvulsant; antibacterial; antifungal; antiparasitic;  
KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;  
KW neurological disease; infection; human; secreted protein.

OS Homo sapiens.

XX  
PN  
WO200190304-A2.

29-NOV-2001.

18-MAY-2001; 2001WO-US016450.

AA  
PR 19-MAY-2000; 2000US-0205515P.

PA (HUMA-) HUMAN GENOME SCI INC.

PI Birse CE, Rosen CA;

WPI: 2002-122018/16.

DR N-PSDB; ABL89642.

Novel 1405 isolated polypeptides, useful for diagnosis, treatment and prevention of neural, immune system, muscular, reproductive, gastrointestinal, pulmonary, cardiovascular, renal and proliferative disorders.

PS	Claim 11; SEQ ID NO 1609; 2081pp + Sequence Listing: English.
XX	

The invention relates to novel genes (ABL89449-ABL90853) and proteins (ABB9040-ABB9044) useful for preventing, treating or ameliorating medical conditions e.g. by protein or gene therapy. The genes are isolated from a range of human tissues disclosed in the specification. The nucleic acids, proteins, antibodies and (ant)agonists are useful in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and ovarian cancer and other cancers of the adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune disorders e.g. Addison's disease, allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c) cardiovascular disorders such as myocardial ischaemias; (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f) infectious diseases such as viral, bacterial, fungal and parasitic infections. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at [ftp.wipo.int/pub/published/pct\\_sequences](http://ftp.wipo.int/pub/published/pct_sequences)

SQ Sequence 261 AA;

### Query Match

Query Match 21.9%; Score 208; DB 5; Length 261;

Best Local Similarity 100.0%; Pred. No. 1.7e-197;

Qy	653	MKLRKILKEKVEAAAGSTSEKDPBEEFNILFHILRLRVEPLLKIRSAGQKVQCQCYFYQI	FM 712
Db	1	MKLRKILKEKVEAAAGSTSEKDPBEEFNILFHILRLRVEPLLKIRSAGQKVQCQCYFYQI	FM 60
Qy	713	EKNEKVGVPITTOQLLEWSFNSNLKFAEAPSCIIIQMPRGKDFKLFKKIPFSL	INTD 772
Db	61	EKNEKVGVPITTOQLLEWSFNSNLKFAEAPSCIIIQMPRGKDFKLFKKIPFSL	INTD 120
Qy	773	LLEDTPRQCRICGGIAMEYECREYDDPDISAGKIKQFCKTNTQVHLHPKRLNHNKYN	PVPS 832
Db	121	LLEDTPRQCRICGGIAMEYECREYDDPDISAGKIKQFCKTNTQVHLHPKRLNHNKYN	PVPS 180
Qy	833	LPKDLPDWDRHGCIPCONMELFAVLCI	860
Db	181	LPKDLPDWDRHGCIPCONMELFAVLCI	208

Search completed: April 18, 2005, 09:39:19

```

search completed:
Job time : 79 secs

```

PR 01-SEP-2000; 2000US-0229343P.  
 PR 01-SEP-2000; 2000US-0229344P.  
 PR 01-SEP-2000; 2000US-0229345P.  
 PR 05-SEP-2000; 2000US-0229509P.  
 PR 05-SEP-2000; 2000US-0229513P.  
 PR 06-SEP-2000; 2000US-0230437P.  
 PR 06-SEP-2000; 2000US-0230438P.  
 PR 08-SEP-2000; 2000US-0231242P.  
 PR 08-SEP-2000; 2000US-0231243P.  
 PR 08-SEP-2000; 2000US-0231244P.  
 PR 08-SEP-2000; 2000US-0231413P.  
 PR 08-SEP-2000; 2000US-0231414P.  
 PR 08-SEP-2000; 2000US-0232080P.  
 PR 12-SEP-2000; 2000US-0232081P.  
 PR 12-SEP-2000; 2000US-0231968P.  
 PR 14-SEP-2000; 2000US-0232397P.  
 PR 14-SEP-2000; 2000US-0232398P.  
 PR 14-SEP-2000; 2000US-0232399P.  
 PR 14-SEP-2000; 2000US-0232400P.  
 PR 14-SEP-2000; 2000US-0232401P.  
 PR 14-SEP-2000; 2000US-0233063P.  
 PR 14-SEP-2000; 2000US-0233064P.  
 PR 14-SEP-2000; 2000US-0233065P.  
 PR 21-SEP-2000; 2000US-0234223P.  
 PR 21-SEP-2000; 2000US-0234224P.  
 PR 25-SEP-2000; 2000US-0234997P.  
 PR 25-SEP-2000; 2000US-0234998P.  
 PR 26-SEP-2000; 2000US-0235484P.  
 PR 27-SEP-2000; 2000US-0235634P.  
 PR 27-SEP-2000; 2000US-0235635P.  
 PR 29-SEP-2000; 2000US-0236327P.  
 PR 29-SEP-2000; 2000US-0236328P.  
 PR 29-SEP-2000; 2000US-0236369P.  
 PR 29-SEP-2000; 2000US-0236369P.  
 PR 29-SEP-2000; 2000US-0236370P.  
 PR 02-OCT-2000; 2000US-0236802P.  
 PR 02-OCT-2000; 2000US-0237032P.  
 PR 02-OCT-2000; 2000US-0237033P.  
 PR 02-OCT-2000; 2000US-0237039P.  
 PR 02-OCT-2000; 2000US-0237040P.  
 PR 13-OCT-2000; 2000US-0239935P.  
 PR 13-OCT-2000; 2000US-0239937P.  
 PR 20-OCT-2000; 2000US-0240960P.  
 PR 20-OCT-2000; 2000US-0241221P.  
 PR 20-OCT-2000; 2000US-0241785P.  
 PR 20-OCT-2000; 2000US-0241786P.  
 PR 20-OCT-2000; 2000US-0241787P.  
 PR 20-OCT-2000; 2000US-0241808P.  
 PR 20-OCT-2000; 2000US-0241809P.  
 PR 01-NOV-2000; 2000US-0241826P.  
 PR 01-NOV-2000; 2000US-0244617P.  
 PR 08-NOV-2000; 2000US-024474P.  
 PR 08-NOV-2000; 2000US-024475P.  
 PR 08-NOV-2000; 2000US-0246476P.  
 PR 08-NOV-2000; 2000US-0246477P.  
 PR 08-NOV-2000; 2000US-0246478P.  
 PR 08-NOV-2000; 2000US-0246523P.  
 PR 08-NOV-2000; 2000US-0246524P.  
 PR 08-NOV-2000; 2000US-0246525P.  
 PR 08-NOV-2000; 2000US-0246526P.  
 PR 08-NOV-2000; 2000US-0246527P.  
 PR 08-NOV-2000; 2000US-0246528P.  
 PR 08-NOV-2000; 2000US-0246532P.  
 PR 08-NOV-2000; 2000US-0246609P.  
 PR 08-NOV-2000; 2000US-0246610P.  
 PR 08-NOV-2000; 2000US-0246611P.  
 PR 08-NOV-2000; 2000US-0246613P.  
 PR 17-NOV-2000; 2000US-0249207P.  
 PR 17-NOV-2000; 2000US-0249208P.  
 PR 17-NOV-2000; 2000US-0249209P.  
 PR 17-NOV-2000; 2000US-0249210P.  
 PR 17-NOV-2000; 2000US-0249211P.  
 PR 17-NOV-2000; 2000US-0249212P.

PR 17-NOV-2000; 2000US-0249213P.  
 PR 17-NOV-2000; 2000US-0249214P.  
 PR 17-NOV-2000; 2000US-0249215P.  
 PR 17-NOV-2000; 2000US-0249216P.  
 PR 17-NOV-2000; 2000US-0249217P.  
 PR 17-NOV-2000; 2000US-0249218P.  
 PR 17-NOV-2000; 2000US-0249244P.  
 PR 17-NOV-2000; 2000US-0249245P.  
 PR 17-NOV-2000; 2000US-0249264P.  
 PR 17-NOV-2000; 2000US-0249265P.  
 PR 17-NOV-2000; 2000US-0249297P.  
 PR 17-NOV-2000; 2000US-0249299P.  
 PR 17-NOV-2000; 2000US-0249300P.  
 PR 01-DEC-2000; 2000US-0250160P.  
 PR 01-DEC-2000; 2000US-0250391P.  
 PR 05-DEC-2000; 2000US-0251030P.  
 PR 05-DEC-2000; 2000US-0251988P.  
 PR 05-DEC-2000; 2000US-0256719P.  
 PR 06-DEC-2000; 2000US-0251479P.  
 PR 08-DEC-2000; 2000US-0251856P.  
 PR 08-DEC-2000; 2000US-0251868P.  
 PR 08-DEC-2000; 2000US-0251869P.  
 PR 08-DEC-2000; 2000US-0251989P.  
 PR 08-DEC-2000; 2000US-0251990P.  
 PR 11-DEC-2000; 2000US-0254097P.  
 PR 05-JAN-2001; 2001US-0259678P.

(HUMA-) HUMAN GENOME SCI INC.

PI Rosen CA, Barash SC, Ruben SM;

XX WPI; 2001-465566/50.

DR N-PSDB; AAS41081.

XX Novel polypeptides and polynucleotides useful for diagnosing, preventing, treating neural, immune system, muscular, reproductive, pulmonary, cardiovascular, renal, proliferative disorders and cancerous diseases.

PS Claim 11; SEQ ID NO 1207; 1180pp; English.

XX The present invention relates to the isolation of novel human enzyme polypeptides, and the cDNA (AAS40795-AAS41684) and genomic sequences encoding them. The enzyme polypeptides of the invention may comprise the functional classes of oxidoreductases, transferases, hydrolases, lyases, isomerases or ligases. The sequences of the invention are useful in the diagnosis, treatment, prevention and/or prognosis of a wide range of disorders including hyperproliferative disorders (e.g. cancer), immunodeficiency disorders (e.g. AIDS) autoimmune disorders (e.g. arthritis), neurological disorders (e.g. Alzheimer's disease), metabolic disorders (e.g. phenylketonuria), inflammatory disorders (e.g. asthma), cardiovascular disorders (e.g. atherosclerosis), blood-related disorders (e.g. haemophilia), reproductive disorders (e.g. infertility) and infectious disorders (e.g. influenza). The polynucleotides of the invention can also be used in gene therapy. AAU22915-AAU23814 represent the novel human enzyme polypeptides of the invention. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 558 AA;

Query Match 23.9%; Score 227; DB 4; Length 558;

Best Local Similarity 100.0%; Pred. No. 4.2e-216; Indels 0; Gaps 0; Matches 227; Conservative 0; Mismatches 0;

QY 344 RSELFYTLNGSSVDSQPSKSNXTWYIDEVAEDPAKSLTEISTDFDRSSPPLQPPVNSL 403

Db 325 RSELFYTLNGSSVDSQPSKSNXTWYIDEVAEDPAKSLTEISTDFDRSSPPLQPPVNSL 384

QY 404 TTENRPHSLPFLSLTKMPTNGSIGHSPLSLSAQSVMEELNTPAQSSPPLAMPFGNSHGL 463

Db 385 TTENRPHSLPFLSLTKMPTNGSIGHSPLSLSAQSVMEELNTPAQSSPPLAMPFGNSHGL 444

Human; primer; detection; diagnosis; antisense therapy; gene therapy.

Homo sapiens.

EP1074617-A2.

07-FEB-2001.

28-JUL-2000; 2000EP-00116126.

29-JUL-1999; 99JP-00248036.

27-AUG-1999; 99JP-00300253.

11-JAN-2000; 2000JP-00118776.

02-MAY-2000; 2000JP-00183767.

09-JUN-2000; 2000JP-00241899.

XX (HELI-) HELIX RES INST.

XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

XX WPI; 2001-318749/34.

DR Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs.

XX Claim 8; SEQ ID NO 18587; 2537pp + Sequence Listing; English.

XX The present invention describes primer sets for synthesizing 5602 full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dT primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesizing polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AAH03166 to AAH13628 and AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632 represent polynucleotides, all of which are used in the exemplification of the present invention

XX Sequence 476 AA;

Query Match 34.8%; Score 330; DB 4; Length 476;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 330; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 344 RSELYTLNGSSVDSQPSKSKNTWYIDEVAEDPAKSLTEISTDFDRSSPPLQPPVNSL 403

Db 126 RSELYTLNGSSVDSQPSKSKNTWYIDEVAEDPAKSLTEISTDFDRSSPPLQPPVNSL 185

QY 404 TTNRFHSLPFLTKMPTNTNGSIGHSPLSLSAQSVMEELNTPAQESPLAMPNGNSHGL 463

Db 186 TTNRFHSLPFLTKMPTNTNGSIGHSPLSLSAQSVMEELNTPAQESPLAMPNGNSHGL 245

QY 464 EVGSLAEVKENPPYGVIRWIGQPGNEVLNAGLELEDECACTDGTGRGTRYFTCALKK 523

Db 246 EVGSLAEVKENPPYGVIRWIGQPGNEVLNAGLELEDECACTDGTGRGTRYFTCALKK 305

QY 524 ALFVKLSKCRPDRFASIQPVSNQIERCNSLAFGGYLSVVVENTPPKMEKGLGIMICK 583

Db 306 ALFVKLSKCRPDRFASIQPVSNQIERCNSLAFGGYLSVVVENTPPKMEKGLGIMICK 365

QY 584 KKGIOGHYNSCYLDSTLFCLEAFSSVLDTVLIRPKKNDVEYYSETQELLRTTEIVNPLRI 643

Db 366 KKGIOGHYNSCYLDSTLFCLEAFSSVLDTVLIRPKKNDVEYYSETQELLRTTEIVNPLRI 425

QY 644 YGYVCATKIMKLRKILEKVEAASGFTSBEK 673

Db 426 YGYVCATKIMKLRKILEKVEAASGFTSBEK 455

RESULT 14

AAU23211

ID AAU23211 standard; protein; 558 AA.

XX AC AAU23211;

XX DT 18-DEC-2001 (first entry)

XX DE Novel human enzyme polypeptide #297.

XX KW Human; oxidoreductase enzyme; transferase; hydrolase; lyase; isomerase; ligase; hyperproliferative disorder; immunodeficiency disorder; autoimmune disorder; neurological disorder; metabolic disorder; inflammatory disorder; cardiovascular disorder; reproductive disorder; blood-related disorder; infectious disorder; cytostatic; anti arthritic; nephrotropic; anticoagulant.

XX OS Homo sapiens.

XX PN WO200155301-A2.

XX PD 02-AUG-2001.

XX PF 17-JAN-2001; 2001WO-US001239.

XX PR 31-JAN-2000; 2000US-0179065P.

PR 04-FEB-2000; 2000US-0180628P.

PR 24-FEB-2000; 2000US-0184664P.

PR 02-MAR-2000; 2000US-0186350P.

PR 16-MAR-2000; 2000US-0189874P.

PR 17-MAR-2000; 2000US-0190076P.

PR 18-APR-2000; 2000US-0198123P.

PR 19-MAY-2000; 2000US-0205515P.

PR 07-JUN-2000; 2000US-0209467P.

PR 28-JUN-2000; 2000US-0214886P.

PR 30-JUN-2000; 2000US-0215135P.

PR 07-JUL-2000; 2000US-0216647P.

PR 07-JUL-2000; 2000US-0216880P.

PR 11-JUL-2000; 2000US-0217487P.

PR 11-JUL-2000; 2000US-0217496P.

PR 14-JUL-2000; 2000US-0218290P.

PR 26-JUL-2000; 2000US-0220963P.

PR 26-JUL-2000; 2000US-0220964P.

PR 14-AUG-2000; 2000US-0224518P.

PR 14-AUG-2000; 2000US-0224519P.

PR 14-AUG-2000; 2000US-0225213P.

PR 14-AUG-2000; 2000US-0225214P.

PR 14-AUG-2000; 2000US-0225266P.

PR 14-AUG-2000; 2000US-0225267P.

PR 14-AUG-2000; 2000US-0225268P.

PR 14-AUG-2000; 2000US-0225270P.

PR 14-AUG-2000; 2000US-0225447P.

PR 14-AUG-2000; 2000US-0225757P.

PR 14-AUG-2000; 2000US-0225758P.

PR 14-AUG-2000; 2000US-0225759P.

PR 18-AUG-2000; 2000US-0226279P.

PR 22-AUG-2000; 2000US-0226681P.

PR 22-AUG-2000; 2000US-0226868P.

PR 22-AUG-2000; 2000US-0227182P.

PR 23-AUG-2000; 2000US-0227009P.

PR 30-AUG-2000; 2000US-0228924P.

PR 01-SEP-2000; 2000US-0229287P.



PA (ASAH-) ASahi KASEI PHARMA CORP.  
XX Matsuda A, Yoneta S;  
XX WPI; 2004-593134/57.  
DR N-PSDB; ADQ95915.  
XX  
XX New purified protein involved in T cell activation, useful for  
PT diagnosing, preventing and/or treating acquired immunodeficiency  
PT syndrome, autoimmune (e.g. rheumatoid arthritis, and diabetes), allergic  
PT and infectious diseases.  
XX  
XX Claim 1; SEQ ID NO 94; 2828pp; English.  
XX  
XX The invention relates to purified proteins and genes encoding them, that  
XX are involved in T cell activation (I) and has an amino acid deletion,  
XX substitution or addition in the amino acid sequences. The methods and  
XX compositions of the present invention are useful for the diagnosis,  
XX prevention and/or treatment of autoimmune disease (rheumatoid arthritis,  
XX asthma, multiple sclerosis and diabetes), allergic disease, infectious  
XX disease, AIDS, and acute or chronic rejection at organ transplant or bone  
XX -marrow transplant. This sequence corresponds to a protein involved in T  
XX cell activation.  
SQ Sequence 953 AA;  
  
Query Match 53.2%; Score 505; DB 8; Length 953;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 605; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
QY 344 RSELYTLNGSSVDSQPSQSKNTWYIDVADPAKSLTEISTDFDRSSPPLQPPVNSL 403  
Db 348 RSELYTLNGSSVDSQPSQSKNTWYIDVADPAKSLTEISTDFDRSSPPLQPPVNSL 407  
  
QY 404 TTNRFHSLPFLTKMPNTNGSIGHSPLSLSAQSVMEELNTAPVQESPLAMPNGSHGL 463  
Db 408 TTNRFHSLPFLTKMPNTNGSIGHSPLSLSAQSVMEELNTAPVQESPLAMPNGSHGL 467  
  
QY 464 EVGSLAEVKNPPFYGVIRWTGQPPGLNEVLAGELEDECACTGTGTRGTRYFTCALKK 523  
Db 468 EVGSLAEVKNPPFYGVIRWTGQPPGLNEVLAGELEDECACTGTGTRGTRYFTCALKK 527  
  
QY 524 ALFVKLSKRDSPASIPQVSNQIERCNSLAFGGLSEVVEENTPPKMEKEGLEIMIGK 583  
Db 528 ALFVKLSKRDSPASIPQVSNQIERCNSLAFGGLSEVVEENTPPKMEKEGLEIMIGK 587  
  
QY 584 KKGIOGHVNSCYLDTLFCFAFSVSLDTVLLRPKEKNDVEYSETQELLRTIYNPLRI 643  
Db 588 KKGIOGHVNSCYLDTLFCFAFSVSLDTVLLRPKEKNDVEYSETQELLRTIYNPLRI 647  
  
QY 644 YGYVCATKIMKILKVEAASGFTSEKQPEEFLNLIFHILRVPEPLKIRSAQKQV 703  
Db 648 YGYVCATKIMKILKVEAASGFTSEKQPEEFLNLIFHILRVPEPLKIRSAQKQV 707  
  
QY 704 DCYFYQIFMEKNEKVGPTIQQLEWSPFNLSLFAEAPSLIITQMPFGKDFKLFKIF 763  
Db 708 DCYFYQIFMEKNEKVGPTIQQLEWSPFNLSLFAEAPSLIITQMPFGKDFKLFKIF 767  
  
QY 764 PSLELNTIDLEDTFPCRIICGLAMVECRICYDDPDISAGKIKQFCIKTQVHLHPR 823  
Db 768 PSLELNTIDLEDTFPCRIICGLAMVECRICYDDPDISAGKIKQFCIKTQVHLHPR 827  
  
QY 824 LNHKNVPSLPKLPDMWRHGCIPQNMELFVLCIETSHYAFVKYKGKDSAWLFFDS 883  
Db 828 LNHKNVPSLPKLPDMWRHGCIPQNMELFVLCIETSHYAFVKYKGKDSAWLFFDS 887  
  
QY 884 MADPDGQGNFIQVTPCPVEGVLYKMSLEDLSLDSRRIOGCARRLLCDAYMCWQSP 943  
Db 888 MADPDGQGNFIQVTPCPVEGVLYKMSLEDLSLDSRRIOGCARRLLCDAYMCWQSP 947  
  
QY 944 TMSLYK 949  
Db 948 TMSLYK 953

RESULT 11  
ADQ95920  
ID ADQ95920 standard; protein; 956 AA.  
XX  
XX AC ADQ95920;  
XX  
XX 07-OCT-2004 (first entry)  
DT  
XX T cell activation associated protein #49.  
DE  
XX  
XX antiallergic; antiarthritic; antiasthmatic; antidiabetic; anti-HIV;  
KW antimicrobial; antirheumatic; immunosuppressive; neuroprotective;  
KW gene therapy; T cell activation; diagnosis; autoimmune disease;  
KW rheumatoid arthritis; asthma; multiple sclerosis; diabetes;  
KW allergic disease; infectious disease; AIDS; chronic rejection; organ;  
KW bone-marrow transplant.  
XX  
XX Homo sapiens.  
OS  
XX WO2004058805-A2.  
PN  
XX 15-JUL-2004.  
XX  
XX 25-DEC-2003; 2003WO-JP016715.  
PF  
XX 26-DEC-2002; 2002JP-00376365.  
PR  
XX 27-DEC-2002; 2002US-0436473P.  
PR  
XX 25-APR-2003; 2003JP-00122113.  
PR  
XX 28-APR-2003; 2003US-0465792P.  
PR  
XX 21-OCT-2003; 2003JP-00360559.  
PR  
XX 22-OCT-2003; 2003US-0512846P.  
XX  
XX (ASAH-) ASahi KASEI PHARMA CORP.  
PA  
XX Matsuda A, Yoneta S;  
PI  
XX WPI; 2004-593134/57.  
DR  
XX N-PSDB; ADQ95919.  
DR  
XX  
XX New purified protein involved in T cell activation, useful for  
PT diagnosing, preventing and/or treating acquired immunodeficiency  
PT syndrome, autoimmune (e.g. rheumatoid arthritis, and diabetes), allergic  
PT and infectious diseases.  
XX  
XX Claim 1; SEQ ID NO 98; 2828pp; English.  
PS  
XX  
XX The invention relates to purified proteins and genes encoding them, that  
XX are involved in T cell activation (I) and has an amino acid deletion,  
XX substitution or addition in the amino acid sequences. The methods and  
XX compositions of the present invention are useful for the diagnosis,  
XX prevention and/or treatment of autoimmune disease (rheumatoid arthritis,  
XX asthma, multiple sclerosis and diabetes), allergic disease, infectious  
XX disease, AIDS, and acute or chronic rejection at organ transplant or bone  
XX -marrow transplant. This sequence corresponds to a protein involved in T  
XX cell activation.  
SQ Sequence 956 AA;  
  
Query Match 53.2%; Score 505; DB 8; Length 956;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 605; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
QY 344 RSELYTLNGSSVDSQPSQSKNTWYIDVADPAKSLTEISTDFDRSSPPLQPPVNSL 403  
Db 351 RSELYTLNGSSVDSQPSQSKNTWYIDVADPAKSLTEISTDFDRSSPPLQPPVNSL 410  
  
QY 404 TTNRFHSLPFLTKMPNTNGSIGHSPLSLSAQSVMEELNTAPVQESPLAMPNGSHGL 463  
Db 411 TTNRFHSLPFLTKMPNTNGSIGHSPLSLSAQSVMEELNTAPVQESPLAMPNGSHGL 470  
  
QY 464 EVGSLAEVKNPPFYGVIRWTGQPPGLNEVLAGELEDECACTGTGTRGTRYFTCALKK 523



RESULT 9	
AD AAB95828	standard; protein; 731 AA.
XX	
AC AAB95828;	
XX	
DT 26-JUN-2001	(first entry)
XX	
DE Human protein sequence SEQ ID NO:18843.	
XX	
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy.	
XX	
OS Homo sapiens.	
XX	
PN EP1074617-A2.	
XX	
PD 07-FEB-2001.	
XX	
PF 28-JUL-2000; 2000EP-00116126.	
XX	
PR 29-JUL-1999; 99JP-00248036.	
PR	
PR 27-AUG-1999; 99JP-00300253.	
PR	
PR 11-JAN-2000; 2000JP-00187776.	
PR	
PR 02-MAY-2000; 2000JP-00183767.	
PR	
PR 09-JUN-2000; 2000JP-00241899.	
XX	
PA (HELI-) HELIX RES INST.	
XX	
PI Ota T, Isegai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;	
PI	
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;	
XX	
DR WPI; 2001-318749/34.	
XX	
PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs.	
XX	
PS Claim 8; SEQ ID NO 18843; 2537pp + Sequence Listing; English.	
XX	
CC The present invention describes primer sets for synthesizing 5602 full-length cDNAs defined in the specification. Where a primer set comprises:	
CC	
CC (a) an oligo-dT primer and an oligonucleotide complementary to the	
CC	
CC complementary strand of a polynucleotide which comprises one of the 5602	
CC	
CC nucleotide sequences defined in the specification, where the	
CC	
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination	
CC	
CC of an oligonucleotide comprising a sequence complementary to the	
CC	
CC complementary strand of a polynucleotide which comprises a 5'-end	
CC	
CC sequence and an oligonucleotide comprising a sequence complementary to a	
CC	
CC polynucleotide which comprises a 3'-end sequence, where the	
CC	
CC oligonucleotide comprises at least 15 nucleotides and the combination of	
CC	
CC the 5'-end sequence/3'-end sequence is selected from those defined in the	
CC	
CC specification. The primer sets can be used in antisense therapy and in	
CC	
CC gene therapy. The primers are useful for synthesizing polynucleotides,	
CC	
CC particularly full-length cDNAs. The primers are also useful for the	
CC	
CC detection and/or diagnosis of the abnormality of the proteins encoded by	
CC	
CC the full-length cDNAs. The primers allow obtaining of the full-length	
CC	
CC cDNAs easily without any special methods. AAH03166 to AAH13628 and	
CC	
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893	
CC	
CC represent human amino acid sequences; and AAH13629 to AAH13632 represent	
CC	
CC oligonucleotides, all of which are used in the exemplification of the	
CC	
CC present invention	
XX	
SQ Sequence 731 AA;	
XX	
Query Match	53.2%; Score 505; DB 4; Length 731;
Best Local Similarity	99.8%; Pred. No. 0;
Matches	605; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 344 RSELYTLNGSSVDSQPSKNTWYIDEVAEDPAKSLTEISTDFDRSSPPLQPPVNSL 403	
Db 126 RSELYTLNGSSVDSQPSKNTWYIDEVAEDPAKSLTEISTDFDRSSPPLQPPVNSL 185	

Qy 404	TTENRFLSLTPSLTKMPNTNGSIHSPLSLSAQSVMEELNTAPVQESPPPLAMPNGSHGL 463
Db	186 TTENRFLSLTPSLTKMPNTNGSIHSPLSLSAQSVMEELNTAPVQESPPPLAMPNGSHGL 245
Qy 464	EVGSLAEVKENPPFYGVIRWICQPGLEVLAGLEDEACAGCTDGTFRGTRYFTCALKK 523
Db	246 EVGSLAEVKENPPFYGVIRWICQPGLEVLAGLEDEACAGCTDGTFRGTRYFTCALKK 305
Qy 524	ALFVKLKSCRDPDSRFASLPQVSNQIERNCSLAFGGVLSVVEENTPPKMEKEGLEIMIGK 583
Db	306 ALFVKLKSCRDPDSRFASLPQVSNQIERNCSLAFGGVLSVVEENTPPKMEKEGLEIMIGK 365
Qy 584	KKGIQGHYNSCYLDSTLFCFAFSSVLDTVLLRPKEKNDVEYYSQTELLRTEIVNPLRI 643
Db	366 KKGIIQGHYNSCYLDSTLFCFAFSSVLDTVLLRPKEKNDVEYYSQTELLRTEIVNPLRI 425
Qy 644	YGYVCATKIMKRLKLEKVEAASGFTSEKDPPEEFLNLFHILRVEPLLKIRSAGQKQV 703
Db	426 YGYVCATKIMKRLKLEKVEAASGFTSEKDPPEEFLNLFHILRVEPLLKIRSAGQKQV 485
Qy 704	DCYFYQIFMEKNEKVGVPITIQOLLEWSFINSNLKFAEAPSCCLIOMPRFGKDFKLPKKIF 763
Db	486 DCYFYQIFMEKNEKVGVPITIQOLLEWSFINSNLKFAEAPSCCLIOMPRFGKDFKLPKKIF 545
Qy 764	PSLELNITDLEDDTPRQCRIICGLAMYECCYDPPDISAGKIKQFCKTCNTQVHLHPKR 823
Db	546 PSLELNITDLEDDTPRQCRIICGLAMYECCYDPPDISAGKIKQFCKTCNTQVHLHPKR 605
Qy 824	LNHNKPNVSLPKDLPDWRHGCIPQNNMELFAVICIETSHYVAFVKYKDDSAMLPFDS 883
Db	606 LNHNKPNVSLPKDLPDWRHGCIPQNNMELFAVICIETSHYVAFVKYKDDSAMLPFDS 665
Qy 884	MADRRGGGNGENIPQVTPCPEVGEYVKMSLEDLHSLDSRRIQGCARRLLCDAYMCMYQSP 943
Db	666 MADRRGGGNGENIPQVTPCPEVGEYVKMSLEDLHSLDSRRIQGCARRLLCDAYMCMYQSP 725
Qy 944	TMSLYK 949
Db	726 TMSLYK 731
RESULT 10	
ADQ95916	
ID	ADQ95916 standard; protein; 953 AA.
XX	
AC	ADQ95916;
XX	
DT	07-OCT-2004 (first entry)
XX	
DE	T cell activation associated protein #47.
XX	
KW	antiallergic; antiarthritic; antiasthmatic; antidiabetic; anti-HIV;
KW	antimicrobial; antirheumatic; immunosuppressive; neuroprotective;
KW	gene therapy; T cell activation; diagnosis; autoimmune disease;
KW	rheumatoid arthritis; asthma; multiple sclerosis; diabetes;
KW	allergic disease; infectious disease; AIDS; chronic rejection; organ;
KW	bone-marrow transplant.
OS	Homo sapiens.
XX	
PN	WO2004058805-A2.
XX	
PD	15-JUL-2004.
XX	
PF	25-DEC-2003; 2003WO-JP016715.
XX	
PR	26-DEC-2002; 2002JP-00376365.
PR	27-DEC-2002; 2002US-0436473P.
PR	25-APR-2003; 2003JP-00122113.
PR	28-APR-2003; 2003US-0465792P.
PR	21-OCT-2003; 2003JP-00360559.
PR	22-OCT-2003; 2003US-0512846P.
XX	

```
Qy 464 EVGSLAEVKNPPFFYGVIRWIGQPPGLNEVLAGELEDEACAGTGTGTGRTGYFTCALKK 523
Dy 471 EVGSLAEVKNPPFFYGVIRWIGQPPGLNEVLAGELEDEACAGTGTGTGRTGYFTCALKK 530
Qy 524 ALFVKLKSRCRPSDRFASLPQVSNQIERCNSLAFGGYLVSEVVENTPPKMEKEGLEIMIGK 583
Dy 531 ALFVKLKSRCRPSDRFASLPQVSNQIERCNSLAFGGYLVSEVVENTPPKMEKEGLEIMIGK 590
Qy 584 KKGIOGHVNSCYLDSTLFCFAFSSVLDTVLLRPKEKNDVEYYSQELLRTTEIVNPLRI 643
Dy 591 KKGIOGHVNSCYLDSTLFCFAFSSVLDTVLLRPKEKNDVEYYSQELLRTTEIVNPLRI 650
Qy 644 YGYVCATKIMKRLKILEKVEAASGFTSEKOPPEEFNLILFHHILRVEPLLKIRSGAQKVQ 703
Dy 651 YGYVCATKIMKRLKILEKVEAASGFTSEKOPPEEFNLILFHHILRVEPLLKIRSGAQKVQ 710
Qy 704 DCYFYQIPEKNEKVGVTIQQLEWSPFNSNLKFAEAPSCILIIQMPRFGKDFLKKIF 763
Dy 711 DCYFYQIPEKNEKVGVTIQQLEWSPFNSNLKFAEAPSCILIIQMPRFGKDFLKKIF 770
Qy 764 PSLELNIITDLEDTPRCRI CGGLAMYECCYDDPDISAGKIKQFCCKTQNTQVHLHPKR 823
Dy 771 PSLELNIITDLEDTPRCRI CGGLAMYECCYDDPDISAGKIKQFCCKTQNTQVHLHPKR 830
Qy 824 LNHNKYNPVSLLPKDLPDWDWRHGCIPQNNMELFAVLCIETSHYVAFVKYKDDSAWLFDD 883
Dy 831 LNHNKYNPVSLLPKDLPDWDWRHGCIPQNNMELFAVLCIETSHYVAFVKYKDDSAWLFDD 890
Qy 884 MADRDGGQGNFNIPOVTPCPEVGEYKMSLEDLHSLDSRRRIQGCARRLLCDAYMCWQSP 943
Dy 891 MADRDGGQGNFNIPOVTPCPEVGEYKMSLEDLHSLDSRRRIQGCARRLLCDAYMCWQSP 950
Qy 944 TMSLYK 949
Dy 951 TMSLYK 956

RESULT 8
ADR99244
ID ADR99244 standard; protein; 960 AA.
AC ADR99244;
XX
DT 02-DEC-2004 (first entry)
DE DKFZp586D1122, SEQ ID 250.
XX
XX Cytostatic; breast cancer; cancer; human; DKFZp586D1122.
OS Homo sapiens.
XX
XX WO2004078035-A2.
PN
XX
XX 16-SEP-2004.
PD
XX
PF 27-FEB-2004; 2004WO-US007268.
PR
XX
XX 28-FEB-2003; 2003US-0450655P.
XX
XX (FARB ) BAYER PHARM CORP.
PA
XX
XX Eveleigh D, Bigwood D;
XX
XX WPI; 2004-653556/63.
DR
XX
XX N-PSDB; ADR99117.
XX
XX Diagnosing breast cancer comprises comparing the level of expression of
PT genes or gene products in a first biological sample taken from a patient
PT with that in a normal patient sample.
XX
XX
XX Claim 3; SEQ ID NO 250; 53pp; English.
XX
```

```
CC The present invention relates to a method (M1) for diagnosing breast
CC cancer in a patient. The method comprises comparing the level of
CC expression of one or more genes or gene products in a biological sample
CC from the patient with that in a normal patient sample, where a difference
CC in the gene expression in the first sample compared to that in the second
CC sample is a diagnostic of the disease. Also claimed are: method (M2) for
CC distinguishing between normal and disease tissues; method (M3) for
CC monitoring the response of a breast cancer patient to treatment with an
CC anti-cancer agent; method (M4) for identifying a compound for treating
CC breast cancer; and an array for distinguishing between normal and disease
CC tissues comprising two or more probes corresponding to genes selected
CC from ADR9895-ADR99121 or comprising two or more polypeptides selected
CC from ADR99122-ADR99248. In M1 and M2 the genes are selected from ADR98955
CC -ADR99121 and the gene products are polypeptides selected from ADR99122-
CC ADR99248. M1 is useful for diagnosing breast cancer. M2 and the array are
CC useful for distinguishing between normal and disease tissue. M3 is useful
CC for monitoring the response of a breast cancer patient to treatment with
CC an anti-cancer agent. M4 is useful for identifying a compound for
CC treating breast cancer. Note: The sequence data for this patent did not
CC form part of the printed specification, but was obtained in electronic
CC format directly from WIFO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 960 AA;
Query Match 63.9%; Score 606; DB 8; Length 960;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 606; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 344 RSELYTTLNGSVDSQPSQSKNTWYIDEVADPAKSLTEISTDFDRSSPPIQPPVNSL 403
Dy 355 RSELYTTLNGSVDSQPSQSKNTWYIDEVADPAKSLTEISTDFDRSSPPIQPPVNSL 414
Qy 404 TTERFHSPLPSLTQMPNTNGSIGHSPISLSAQSWMEELNTPAQVESPPLAMPNGSHGL 463
Dy 415 TTERFHSPLPSLTQMPNTNGSIGHSPISLSAQSWMEELNTPAQVESPPLAMPNGSHGL 474
Qy 464 EVGSLAEVKNPPFFYGVIRWIGQPPGLNEVLAGELEDEACAGTGTGTGRTGYFTCALKK 523
Dy 475 EVGSLAEVKNPPFFYGVIRWIGQPPGLNEVLAGELEDEACAGTGTGTGRTGYFTCALKK 534
Qy 524 ALFVKLKSRCRPSDRFASLPQVSNQIERCNSLAFGGYLVSEVVENTPPKMEKEGLEIMIGK 583
Dy 535 ALFVKLKSRCRPSDRFASLPQVSNQIERCNSLAFGGYLVSEVVENTPPKMEKEGLEIMIGK 594
Qy 584 KKGIOGHVNSCYLDSTLFCFAFSSVLDTVLLRPKEKNDVEYYSQELLRTTEIVNPLRI 643
Dy 595 KKGIOGHVNSCYLDSTLFCFAFSSVLDTVLLRPKEKNDVEYYSQELLRTTEIVNPLRI 654
Qy 644 YGYVCATKIMKRLKILEKVEAASGFTSEKOPPEEFNLILFHHILRVEPLLKIRSGAQKVQ 703
Dy 655 YGYVCATKIMKRLKILEKVEAASGFTSEKOPPEEFNLILFHHILRVEPLLKIRSGAQKVQ 714
Qy 704 DCYFYQIPEKNEKVGVTIQQLEWSPFNSNLKFAEAPSCILIIQMPRFGKDFLKKIF 763
Dy 715 DCYFYQIPEKNEKVGVTIQQLEWSPFNSNLKFAEAPSCILIIQMPRFGKDFLKKIF 774
Qy 764 PSLELNIITDLEDTPRCRI CGGLAMYECCYDDPDISAGKIKQFCCKTQNTQVHLHPKR 823
Dy 775 PSLELNIITDLEDTPRCRI CGGLAMYECCYDDPDISAGKIKQFCCKTQNTQVHLHPKR 834
Qy 824 LNHNKYNPVSLLPKDLPDWDWRHGCIPQNNMELFAVLCIETSHYVAFVKYKDDSAWLFDD 883
Dy 835 LNHNKYNPVSLLPKDLPDWDWRHGCIPQNNMELFAVLCIETSHYVAFVKYKDDSAWLFDD 894
Qy 884 MADRDGGQGNFNIPOVTPCPEVGEYKMSLEDLHSLDSRRRIQGCARRLLCDAYMCWQSP 943
Dy 895 MADRDGGQGNFNIPOVTPCPEVGEYKMSLEDLHSLDSRRRIQGCARRLLCDAYMCWQSP 954
Qy 944 TMSLYK 949
Dy 955 TMSLYK 960
```

CC	sequence represents the human CYLD polypeptide
XX	
SQL	Sequence 956 AA;
	Query Match 63.9%; Score 606; DB 6; Length 956;
	Best Local Similarity 100.0%; Pred. No. 0;
	Matches 606; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	344 RSELYFTLNGSSVDSQPSQSKNTWYIDEVADPAKSLTEISTDFDRSSPPLQPPVNSL 403
DB	351 RSELYFTLNGSSVDSQPSQSKNTWYIDEVADPAKSLTEISTDFDRSSPPLQPPVNSL 410
QY	404 TTEHRSFLPSLTQMPNTNGSIGHSPLSLAQSVMEELNTPAQVESPPPLAMPNGSHGL 463
DB	411 TTEHRSFLPSLTQMPNTNGSIGHSPLSLAQSVMEELNTPAQVESPPPLAMPNGSHGL 470
QY	464 EVGSLAEVKEPPFYGVIRWIGQPPGLNEVLAGELEDEACAGTGTGTRGTYFYFCALKK 523
DB	471 EVGSLAEVKEPPFYGVIRWIGQPPGLNEVLAGELEDEACAGTGTGTRGTYFYFCALKK 530
QY	524 ALFVKLKSCTPDSRPFASLPQVSNQIERCNSLAFGGYLVSEVVENTPPKMEKEGLEIMIGK 583
DB	531 ALFVKLKSCTPDSRPFASLPQVSNQIERCNSLAFGGYLVSEVVENTPPKMEKEGLEIMIGK 590
QY	584 KGIGQHYNSCYLDTLFCFAFSSVLTDLRLPKENDVEYYSQTEQLLRTEIVNPLRI 643
DB	591 KGIGQHYNSCYLDTLFCFAFSSVLTDLRLPKENDVEYYSQTEQLLRTEIVNPLRI 650
QY	644 YGYVCATKIMKRLKILEKVEAASGFTSEKDPPEFLNLFHILRVEPLLKIRSGQKQV 703
DB	651 YGYVCATKIMKRLKILEKVEAASGFTSEKDPPEFLNLFHILRVEPLLKIRSGQKQV 710
QY	704 DCYFYQIFMEKNEKVGVPITIQELLEWSFINSNLKFAEAPSCLIIQMPREGDKFLFKKIF 763
DB	711 DCYFYQIFMEKNEKVGVPITIQELLEWSFINSNLKFAEAPSCLIIQMPREGDKFLFKKIF 770
QY	764 PSLEINITDLEDTQRCRICGLAMYECCYDDPDISAGIKIKQFCKTCNTQVHLHPRK 823
DB	771 PSLEINITDLEDTQRCRICGLAMYECCYDDPDISAGIKIKQFCKTCNTQVHLHPRK 830
QY	824 LNHKYNVSLPKDLPDWRHGCIPQNMELFAVICIETSHYVAFVKYKDDSAWLFFDS 883
DB	831 LNHKYNVSLPKDLPDWRHGCIPQNMELFAVICIETSHYVAFVKYKDDSAWLFFDS 890
QY	884 MADRDGGQGNFIPQVTPCPEGEYLVKMSLEDLSLDSRRIOGCARRLLCDAYMCYQSP 943
DB	891 MADRDGGQGNFIPQVTPCPEGEYLVKMSLEDLSLDSRRIOGCARRLLCDAYMCYQSP 950
QY	944 TMSLYK 949
DB	951 TMSLYK 956
RESULT 7	
ID	ADRI14489
XX	ADRI14489 standard; protein; 956 AA.
AC	
XX	ADRI14489;
DT	21-OCT-2004 (first entry)
XX	
DE	*Human NF-kappaB pathway-associated protein SeqID490.
XX	
KW	NF-kappaB pathway; antiinflammatory; cytostatic; hepatotropic; virucide;
KW	antiarthritic; antiarthritis; gastrointestinal-Gen; antiasthmatic;
KW	antiarteriosclerotic; immunomodulator; cerebroprotective; vasotropic;
KW	immunosuppressive; vulnary; gene therapy; immune disorder;
KW	inflammatory disorder; NF-kappaB regulation; cancer; aberrant apoptosis;
KW	hepatic disorder; Hodgkin's lymphoma; haematopoietic tumour;
KW	hyper-IgM syndrome; hypohidrotic ectodermal dysplasia;
KW	X-linked anhidrotic ectodermal dysplasia; immunodeficiency;
KW	viral infection; HIV-1; HTLV-1; hepatitis B; hepatitis C; EBV; influenza;
KW	viral replication; host cell survival; evasion of immune response;

KW	rheumatoid arthritis; inflammatory bowel disease; colitis; asthma;
KW	atherosclerosis; cachexia; euthyroid sick syndrome; stroke; EAE;
KW	autoimmune disorder; hyper immune activity;
KW	aberrant acute phase response; hypercongenital condition; birth defect;
KW	neurotic lesion; wound; organ transplant rejection;
KW	aberrant signal transduction; proliferating disorder; cancer;
KW	HIV propagation; human.
XX	
OS	Homo sapiens.
XX	
XX	WO2004065577-A2.
XX	
PN	
XX	05-AUG-2004.
PD	
XX	
PF	13-JAN-2004; 2004WO-US000798.
XX	
XX	14-JAN-2003; 2003US-0440068P.
PR	
PR	12-MAY-2003; 2003US-0469757P.
XX	
XX	(BRIM ) BRISTOL-MYERS SQUIBB CO.
PA	
XX	
XX	Nadler SG, Neubauer MG, Feder JN, Carman J;
PI	
XX	
XX	WPI; 2004-562168/54.
DR	N-PSDB; ADRI4488.
XX	
XX	New isolated polynucleotides and polypeptides associated with NF-kappaB pathway, useful for diagnosing, treating, or preventing disorders or diseases associated with NF-kappaB pathway.
PT	
PT	
XX	
PS	Claim 6; SEQ ID NO 490; 237pp; English.
XX	
CC	This invention relates to the novel association of protein sequences (and the genes which encode them) to the NF-kappaB pathway. The invention may be useful for the production of compounds with an antiinflammatory, cytosstatic, hepatotropic, virucide, antiarthritic, antiarthritis, antiarteriosclerotic, gastrointestinal-Gen, antiasthmatic, antiarteriosclerotic, immunomodulator, cerebroprotective, vasotropic, immunosuppressive or vulnary activity or for gene therapy. The proteins and nucleotides are useful for diagnosing, preventing, treating, or ameliorating conditions or diseases associated with the NF-kappaB pathway. The condition is an immune disorder, an inflammatory disorder, an inflammatory disorder related to aberrant NF-kappaB regulation, cancer, aberrant apoptosis, hepatic disorders, Hodgkin's lymphomas, haematopoietic tumours, hyper-IgM syndromes, hypohidrotic ectodermal dysplasia, X-linked anhidrotic ectodermal dysplasia, immunodeficiency, viral infections, HIV-1, hepatitis B, hepatitis C, EBV, influenza, viral replication, host cell survival, evasion of immune responses, rheumatoid arthritis, inflammatory bowel disease, colitis, asthma, atherosclerosis, cachexia, euthyroid sick syndrome, stroke, EAE, autoimmune disorders, disorders related to hyper immune activity, disorders related to aberrant acute phase responses, hypercongenital conditions, birth defects, necrotic lesions, wounds, organ transplant rejection, conditions related to organ transplant rejection, disorders related to aberrant signal transduction, proliferating disorders, cancers and HIV propagation in cells infected with other viruses. The present sequence is that of a human protein which is subject to the novel association with the NF-kappaB pathway of the invention. Note: This sequence does not appear in the specification but was obtained by the indexer from Genbank.
CC	
XX	
SQL	Sequence 956 AA;
	Query Match 63.9%; Score 606; DB 8; Length 956;
	Best Local Similarity 100.0%; Pred. No. 0;
	Matches 606; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	344 RSELYFTLNGSSVDSQPSQSKNTWYIDEVADPAKSLTEISTDFDRSSPPLQPPVNSL 403
DB	351 RSELYFTLNGSSVDSQPSQSKNTWYIDEVADPAKSLTEISTDFDRSSPPLQPPVNSL 410
QY	404 TTEHRSFLPSLTQMPNTNGSIGHSPLSLAQSVMEELNTPAQVESPPPLAMPNGSHGL 463
DB	411 TTEHRSFLPSLTQMPNTNGSIGHSPLSLAQSVMEELNTPAQVESPPPLAMPNGSHGL 470

KW	rheumatoid arthritis; inflammatory bowel disease; colitis; asthma;
KW	atherosclerosis; cachexia; euthyroid sick syndrome; stroke; EAE;
KW	autoimmune disorder; hyper immune activity;
KW	aberrant acute phase response; hypercongenital condition; birth defect;
KW	necrotic lesion; wound; organ transplant rejection;
KW	aberrant signal transduction; proliferating disorder; cancer;
KW	HIV propagation; human.
XX	Homo sapiens.
OS	WO2004065577-A2.
PN	05-AUG-2004.
PD	13-JAN-2004; 2004WO-US000798.
XX	14-JAN-2003; 2003US-0440068P.
PR	12-MAY-2003; 2003US-0469757P.
XX	(BRIM ) BRISTOL-MYERS SQUIBB CO.
PA	Nadler SG, Neubauer MG, Feder JN, Carman J;
XX	WPI; 2004-562168/54.
DR	N-PSDB; ADRI14488.
XX	New isolated polynucleotides and polypeptides associated with NF-kappaB pathway, useful for diagnosing, treating, or preventing disorders or diseases associated with NF-kappaB pathway.
PT	Claim 6; SEQ ID NO 490; 237pp; English.
CC	This invention relates to the novel association of protein sequences (and the genes which encode them) to the NF-kappaB pathway. The invention may be useful for the production of compounds with an antiinflammatory, cytostatic, hepatotropic, virucide, antiarthritic, antiarteriosclerotic, immunomodulator, cerebroprotective, vasotropic, immunosuppressive or vulnary activity or for gene therapy. The proteins and nucleotides are useful for diagnosing, preventing, treating, or ameliorating conditions or diseases associated with the NF-kappaB pathway. The condition is an immune disorder, an inflammatory disorder, an inflammatory disorder related to aberrant NF-kappaB regulation, cancer, aberrant apoptosis, hepatic disorders, Hodgkin's lymphomas, haematopoietic tumours, hyper-IgM syndromes, hypohidrotic ectodermal dysplasia, X-linked anhidrotic ectodermal dysplasia, immunodeficiency, viral infections, HIV-1, HTLV-1, hepatitis B, hepatitis C, EBV, influenza, viral replication, host cell survival, evasion of immune responses, rheumatoid arthritis, inflammatory bowel disease, colitis, asthma, atherosclerosis, cachexia, euthyroid sick syndrome, stroke, EAE, autoimmune disorders, disorders related to hyper immune activity, disorders related to aberrant acute phase responses, hypercongenital conditions, birth defects, necrotic lesions, wounds, organ transplant rejection, conditions related to organ transplant rejection, disorders related to aberrant signal transduction, proliferating disorders, cancers and HIV propagation in cells infected with other viruses. The present sequence is that of a human protein which is subject to the novel association with the NF-kappaB pathway of the invention. Note: This sequence does not appear in the specification but was obtained by the indexer from Genbank.
XX	Sequence 956 AA;
SQL	Query Match 63.9%; Score 606; DB 8; Length 956;
	Best Local Similarity 100.0%; Pred. No. 0;
	Matches 606; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy	344 RSELYFTLNGSSVDSQPSQSKNTWYIDEVADPAKSLTEISTDFDRSSPPLQPPVNSL 403
Db	351 RSELYFTLNGSSVDSQPSQSKNTWYIDEVADPAKSLTEISTDFDRSSPPLQPPVNSL 410
Qy	404 TTEHFSLPSLTQMPNTNGSIGHSPLSLAQSVMEELNTPAQVESPPPLAMPNGSHGL 463
Db	411 TTEHFSLPSLTQMPNTNGSIGHSPLSLAQSVMEELNTPAQVESPPPLAMPNGSHGL 470

07-OCT-2004 (first entry)  
T cell activation associated protein #48.  
antiallergic; antiarthritic; antidiabetic; antidiabetic; anti-HIV;  
antimicrobial; antirheumatic; immunosuppressive; neuroprotective;  
gene therapy; T cell activation; diagnosis; autoimmune disease;  
rheumatoid arthritis; asthma; multiple sclerosis; diabetes;  
allergic disease; infectious disease; AIDS; chronic rejection; organ;  
bone-marrow transplant.  
Homo sapiens.  
WO2004058805-A2.  
15-JUL-2004.  
25-DEC-2003; 2003WO-JP016715.  
26-DEC-2002; 2002JP-00376365.  
27-DEC-2002; 2002US-0436473P.  
25-APR-2003; 2003JP-00122113.  
28-APR-2003; 2003US-0465792P.  
21-OCT-2003; 2003JP-00360559.  
22-OCT-2003; 2003US-0512846P.  
(ASAH-) ASahi KASEI PHARMA CORP.  
Matsuda A, Yoneta S;  
WPI; 2004-593134/57.  
N-PSDB; ADQ95917.  
New purified protein involved in T cell activation, useful for  
diagnosing, preventing and/or treating acquired immunodeficiency  
syndrome, autoimmune (e.g. rheumatoid arthritis, and diabetes), allergic  
and infectious diseases.  
Claim 1; SEQ ID NO 96; 2828pp; English.  
The invention relates to purified proteins and genes encoding them, that  
are involved in T cell activation (I) and has an amino acid deletion,  
substitution or addition in the amino acid sequences. The methods and  
compositions of the present invention are useful for the diagnosis,  
prevention and/or treatment of autoimmune disease (rheumatoid arthritis,  
asthma, multiple sclerosis and diabetes), allergic disease, infectious  
disease, AIDS, and acute or chronic rejection at organ transplant or bone  
-marrow transplant. This sequence corresponds to a protein involved in T  
cell activation.  
Sequence 953 AA;  
Query Match 63.9%; Score 606; DB 8; Length 953;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 606; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 344 RSELYTLNGSSVDSQPSKNTWYIDEVAEDPAKSLTEISTDPDRSSPLQPPVNSL 403  
DB 348 RSELYTLNGSSVDSQPSKNTWYIDEVAEDPAKSLTEISTDPDRSSPLQPPVNSL 407  
QY 404 TTENRHSLPSTLKPNTNGSIGHSPUSLAQSVMEELNAPVOESPLAMPNGSHGL 463  
DB 408 TTENRHSLPSTLKPNTNGSIGHSPUSLAQSVMEELNAPVOESPLAMPNGSHGL 467  
QY 464 EVGSLAEVKEPPFYGVIRWICQPPGLNEVLAGELEDECACTDGTFRGTRYFTCALKK 523  
DB 468 EVGSLAEVKEPPFYGVIRWICQPPGLNEVLAGELEDECACTDGTFRGTRYFTCALKK 527  
QY 524 ALFKLKSCKRPSFASLPQVSNQIERNCSLAFGGYLSVEVVENTPPKMEKEGLEIMIGK 583  
DB 528 ALFKLKSCKRPSFASLPQVSNQIERNCSLAFGGYLSVEVVENTPPKMEKEGLEIMIGK 587  
QY 584 KKGIOGHVNSCYLSTLFLCLFAFSSVLDVLLRPKEKNDVEYYSQTQELLRTIVNPLRI 643

Db 588 KKGIOGHVNSCYLSTLFLCLFAFSSVLDVLLRPKEKNDVEYYSQTQELLRTIVNPLRI 647  
QY 644 YGVYCATKIMKRLKILEKVEAASGFTSEKDEEFLNIPHLRVEPLLKIRSAGQKQV 703  
Db 648 YGVYCATKIMKRLKILEKVEAASGFTSEKDEEFLNIPHLRVEPLLKIRSAGQKQV 707  
QY 704 DCYFYQIFMEKNEKVGVPITIQOLLEWSFINSNLKFAEAPSCSLTIQMPRFGKDFKLPFKKIF 763  
Db 708 DCYFYQIFMEKNEKVGVPITIQOLLEWSFINSNLKFAEAPSCSLTIQMPRFGKDFKLPFKKIF 767  
QY 764 PSLELNITDLEDTTPRCRIQGLLAMEYECREYDDPDISAGIKIQPKCTCNTQVHLHPKR 823  
Db 768 PSLELNITDLEDTTPRCRIQGLLAMEYECREYDDPDISAGIKIQPKCTCNTQVHLHPKR 827  
QY 824 LNHKNYPVSLPKDLPDMDWRHGCIPCONMELFVLCIETSHYVAFVKYKGDPSAWLFFDS 883  
Db 828 LNHKNYPVSLPKDLPDMDWRHGCIPCONMELFVLCIETSHYVAFVKYKGDPSAWLFFDS 887  
QY 884 MADRDGGGNGFNIPQVTPCPEVGEYLKMSLEDLHSLDSRRIQGCARRLLCDAYMCMYQSP 943  
Db 888 MADRDGGGNGFNIPQVTPCPEVGEYLKMSLEDLHSLDSRRIQGCARRLLCDAYMCMYQSP 947  
QY 944 TMSLYK 949  
Db 948 TMSLYK 953

RESULT 6  
ABB82783  
ID ABB82783 standard; protein; 956 AA.  
XX  
AC ABB82783;  
DT 18-MAR-2003 (first entry)  
XX  
DE Human CYLD polypeptide.  
XX  
KW CD40; NF-kappaB essential modulator; NEMO; CYLD; immunosuppressive;  
antiinflammatory; cytostatic; gene therapy; human.  
XX  
OS Homo sapiens.  
XX  
PN WO200292761-A2.  
XX  
PD 21-NOV-2002.  
XX  
PF 07-MAY-2002; 2002WO-US014570.  
XX  
PR 08-MAY-2001; 2001US-00851673.  
XX  
PA (IMMV ) IMMUNEX CORP.  
XX  
PI Derry JMW, Fanslow WC, Dougall WC;  
XX  
XX WPI; 2003-120669/11.  
DR N-PSDB; ABV75394.  
XX  
PT Identifying compounds that alter one or more biological activities of  
CD40 by modulating the binding of NEMO and CYLD, useful for treating  
disorders of the immune system, and inflammatory and cancer diseases.  
XX  
PS Claim 1; Page 38-44; 48pp; English.  
XX  
CC The invention relates to methods of identifying compounds that alter one/  
more biological activities of CD40. One method involves screening for  
molecules that modulate the binding of NF-kappaB essential modulator  
(NEMO) and CYLD. The methods and compositions of the invention of  
determining compounds that agonize or antagonize a CD40 signaling  
activity, are useful for the further definition of CD40-mediated  
signaling pathways, and for manipulation of CD40-mediated cellular  
responses. They also provide therapeutic agents for treating disorders of  
the immune system, and inflammatory and cancer diseases. The present

Db 374 KGIQGHVNSCYLDSTLFCFAFSSVLDTVLLRPKEKNDVEYVSETQELLRTTEIVNPLRI 433  
Qy 644 YGYVCATKIMLKIRKILEKVEAASGFTSEKDPPEEFNLILFHHILRVEPLLKIRSAQKQV 703  
Db 434 YGYVCATKIMLKIRKILEKVEAASGFTSEKDPPEEFNLILFHHILRVEPLLKIRSAQKQV 493  
Qy 704 DCYFYQIFMEKNEKVGVPVPTIQQLLEWSFINSNLKFAEAPSCILIIQMPRFGKDFKLFKKIF 763  
Db 494 DCYFYQIFMEKNEKVGVPVPTIQQLLEWSFINSNLKFAEAPSCILIIQMPRFGKDFKLFKKIF 553  
Qy 764 PSLELNITDLEDTPRQCRICGGLAMYEKRECYDDPDISAGKIKQFCCKTCNTQVHLHPKR 823  
Db 554 PSLELNITDLEDTPRQCRICGGLAMYEKRECYDDPDISAGKIKQFCCKTCNTQVHLHPKR 613  
Qy 824 LNHNKYNPVSPLPKDLPDWDWRHGCIPQNNMELFAVLICETSHYVAFVKYKQDSAWLFFDS 883  
Db 614 LNHNKYNPVSPLPKDLPDWDWRHGCIPQNNMELFAVLICETSHYVAFVKYKQDSAWLFFDS 673  
Qy 884 MADRDGGQNGFNIPQVTPCPEVGEYVKMSLEDLHSLDSRRIOGCARRLLCDAYMCWQSP 943  
Db 674 MADRDGGQNGFNIPQVTPCPEVGEYVKMSLEDLHSLDSRRIOGCARRLLCDAYMCWQSP 733  
Qy 944 TMSLYK 949  
Db 734 TMSLYK 739

RESULT 4  
ADC24816  
ID ADC24816 standard; protein; 953 AA.  
XX AC  
AC ADC24816;  
DT 18-DEC-2003 (first entry)  
XX Human breast specific polypeptide (BSP) DEX0238\_137, SEQ ID NO:137.  
DE Human; breast specific polypeptide; BSP; breast specific nucleic acid;  
XX BSNAs; breast cancer; metastasis; non-cancerous disease; breast tissue;  
KW identification; monitoring; diagnosis;  
KW engineered breast tissue production; transgenic animal; drug screening;  
KW cytostatic; gene therapy; vaccine; chromosome 16p13.3.  
XX Homo sapiens.  
OS  
XX WO2003020900-A2.  
FN  
XX 13-MAR-2003.  
PD  
XX 29-AUG-2002; 2002WO-US027777.  
PF  
XX 31-AUG-2001; 2001US-0316306P.  
PR  
XX (DIAD-) DIADEXUS INC.  
PA  
XX Sun Y, Liu C, Salceda S;  
PI  
XX WPI; 2003-290182/28.  
DR N-PSDB; ADC24898.  
XX  
XX New breast specific polypeptide useful for identifying, diagnosing,  
PT monitoring, staging, imaging and treating breast cancer and non-cancerous  
PT disease states in breast.  
XX  
XX Disclosure; SEQ ID NO 137; 264pp; English.  
PS  
XX The invention relates to breast specific polypeptides (BSPs) and nucleic  
CC acids (breast specific nucleic acids; BSNAs) encoding them. The invention  
CC also relates to vectors and host cells comprising a BSNAs sequence;  
CC antibodies against BSPs; the recombinant production of BSPs; methods of  
CC detection of BSNAs or BSPs in a sample; kits for detecting a risk of  
CC cancer or presence of cancer in a patient; and vaccines comprising a BSNAs  
CC or BSP. The invention additionally discloses fragments, mutants, fusion

CC proteins, homologous proteins and allelic variants of BSPs; methods for  
CC identifying and designing agonists and antagonists of BSPs; methods for  
CC identifying and monitoring breast tissue; producing engineered breast  
CC tissue for treatment and research; producing transgenic animals and cells  
CC comprising BSNAs sequences; aptamers evolved to bind specifically to BSPs;  
CC and single exon probes based on BSNAs sequences. BSPs, BSNAs and  
CC antibodies against BSPs are useful for identifying, diagnosing,  
CC monitoring, staging, imaging and treating breast cancer (including breast  
CC cancer metastases), and non-cancerous disease states in breast tissue.  
CC BSPs and BSNAs may additionally be used to identify and monitor breast  
CC tissue, in screening for BSP agonists and antagonists, and in the  
CC production of engineered breast tissue for treatment or research. BSNAs  
CC may also be used in gene therapy and in the production of transgenic  
CC animals and cells. The present sequence represents a breast specific  
CC polypeptide (BSP) disclosed in the invention.  
XX  
SQ Sequence 953 AA;  
Query Match 63.9%; Score 606; DB 7; Length 953;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 606; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 344 RSELYTLNGSSVDSQPSKSNWTWYIDEVAEDPAKSLTEISTDFDRSSPPLQPPVNSL 403  
Db 348 RSELYTLNGSSVDSQPSKSNWTWYIDEVAEDPAKSLTEISTDFDRSSPPLQPPVNSL 407  
Qy 404 TTENRFHSLPFLSLTKMPTNGSIGHSPISLSAQSVMEELNAPVQESPLAMPNGSHGL 463  
Db 408 TTENRFHSLPFLSLTKMPTNGSIGHSPISLSAQSVMEELNAPVQESPLAMPNGSHGL 467  
Qy 464 EYGSIAEVENPPFYGVIRWIGQPPGLNEVLAGELEDEACAGTGTGTFRTYFTCALKK 523  
Db 468 EYGSIAEVENPPFYGVIRWIGQPPGLNEVLAGELEDEACAGTGTGTFRTYFTCALKK 527  
Qy 524 ALFVKLKSCRPDSRFASLQPVNSQIERCNSLAFGGYLSVEVVENTPPKMEKEGLEIMIGK 583  
Db 528 ALFVKLKSCRPDSRFASLQPVNSQIERCNSLAFGGYLSVEVVENTPPKMEKEGLEIMIGK 587  
Qy 584 KGIQGHVNSCYLDSTLFCFAFSSVLDTVLLRPKEKNDVEYVSETQELLRTTEIVNPLRI 643  
Db 588 KGIQGHVNSCYLDSTLFCFAFSSVLDTVLLRPKEKNDVEYVSETQELLRTTEIVNPLRI 647  
Qy 644 YGYVCATKIMLKIRKILEKVEAASGFTSEKDPPEEFNLILFHHILRVEPLLKIRSAQKQV 703  
Db 648 YGYVCATKIMLKIRKILEKVEAASGFTSEKDPPEEFNLILFHHILRVEPLLKIRSAQKQV 707  
Qy 704 DCYFYQIFMEKNEKVGVPVPTIQQLLEWSFINSNLKFAEAPSCILIIQMPRFGKDFKLFKKIF 763  
Db 708 DCYFYQIFMEKNEKVGVPVPTIQQLLEWSFINSNLKFAEAPSCILIIQMPRFGKDFKLFKKIF 767  
Qy 764 PSLELNITDLEDTPRQCRICGGLAMYEKRECYDDPDISAGKIKQFCCKTCNTQVHLHPKR 823  
Db 768 PSLELNITDLEDTPRQCRICGGLAMYEKRECYDDPDISAGKIKQFCCKTCNTQVHLHPKR 827  
Qy 824 LNHNKYNPVSPLPKDLPDWDWRHGCIPQNNMELFAVLICETSHYVAFVKYKQDSAWLFFDS 883  
Db 828 LNHNKYNPVSPLPKDLPDWDWRHGCIPQNNMELFAVLICETSHYVAFVKYKQDSAWLFFDS 887  
Qy 884 MADRDGGQNGFNIPQVTPCPEVGEYVKMSLEDLHSLDSRRIOGCARRLLCDAYMCWQSP 943  
Db 888 MADRDGGQNGFNIPQVTPCPEVGEYVKMSLEDLHSLDSRRIOGCARRLLCDAYMCWQSP 947  
Qy 944 TMSLYK 949  
Db 948 TMSLYK 953  
RESULT 5  
ADQ95918  
ID ADQ95918 standard; protein; 953 AA.  
XX AC  
AC ADQ95918;  
XX







XX The invention relates to human nucleic acids (AA157798-AA161369) and the  
CC encoded polypeptides (AAM38642-AA42213) with neurotropic,  
CC immunosuppressant and cytostatic activity. The polynucleotides are useful  
CC in gene therapy. A composition containing a polypeptide or polynucleotide  
CC of the invention may be used to treat diseases of the peripheral nervous  
CC system, such as peripheral nervous injuries, peripheral neuropathy and  
CC localised neuropathies and central nervous system diseases, such as  
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
CC utilisation of the activities such as: immune system suppression,  
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
CC assays for receptor activity, arthritis and inflammation, leukaemia and  
CC C.N.S disorders. Note: The sequence data for this patent did not form  
CC part of the printed specification

XX SQ Sequence 685 AA;

Query Match 63.9%; Score 606; DB 4; Length 685;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 606; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 344 RSELYTLNGSSVDSQPSKNTWYIDEVAEDPAKSLTEISTDRSPPLQPPVNSL 403  
DB 80 RSELYTLNGSSVDSQPSKNTWYIDEVAEDPAKSLTEISTDRSPPLQPPVNSL 139  
QY 404 TTEFRHSLPSPKMPNTNGSIGHSPLSLSAQSVMEELNAPVQESPLAMPNSHGL 463  
DB 140 TTEFRHSLPSPKMPNTNGSIGHSPLSLSAQSVMEELNAPVQESPLAMPNSHGL 199  
QY 464 EVGLAEVKNPPFVGIRWVGQPLNEVLAGELEDECACTGTGTRFYTCALKK 523  
DB 200 EVGLAEVKNPPFVGIRWVGQPLNEVLAGELEDECACTGTGTRFYTCALKK 259  
QY 524 ALFVKLSRCDPSFASLPQVSNQIERNCSLAFGGYLSVEVEENTPPKMEKGLIMIGK 583  
DB 260 ALFVKLSRCDPSFASLPQVSNQIERNCSLAFGGYLSVEVEENTPPKMEKGLIMIGK 319  
QY 584 KKGIOGHYNSCYLDSTLFCPLFAFSVLDVLLRPKEKNDVEYSETQELLRTIYNPLRI 643  
DB 320 KKGIOGHYNSCYLDSTLFCPLFAFSVLDVLLRPKEKNDVEYSETQELLRTIYNPLRI 379  
QY 644 YGYVCATKIMKRLKILEKVEASGFTSEKDPPEELNLFPHILRVEPLLKIRSAGQVK 703  
DB 380 YGYVCATKIMKRLKILEKVEASGFTSEKDPPEELNLFPHILRVEPLLKIRSAGQVK 439  
QY 704 DCYFYQIFMEKNEKVGVPITQOLLEWSFINSLKFAEAPSCILIQMPRFGKDFKLFKIF 763  
DB 440 DCYFYQIFMEKNEKVGVPITQOLLEWSFINSLKFAEAPSCILIQMPRFGKDFKLFKIF 499  
QY 764 PSLELNITDLEDTPQRCICGGLAMYECCYDDPDISAGKIKQFCCTNTQVHLHPKR 823  
DB 500 PSLELNITDLEDTPQRCICGGLAMYECCYDDPDISAGKIKQFCCTNTQVHLHPKR 559  
QY 824 LNHHKYNVSLPKDLPDMDWRHGCTPCQNMELFAVLCTIETSHYAPVKYKDDSAWLFPDS 883  
DB 560 LNHHKYNVSLPKDLPDMDWRHGCTPCQNMELFAVLCTIETSHYAPVKYKDDSAWLFPDS 619  
QY 884 MADRDGGONGFNIQVTPCPVGEVYKMSLDLHSLDRRIQGCARRLLCDAYMCMYOSP 943  
DB 620 MADRDGGONGFNIQVTPCPVGEVYKMSLDLHSLDRRIQGCARRLLCDAYMCMYOSP 679  
QY 944 TMSLYK 949  
DB 680 TMSLYK 685

RESULT 2  
AA919154  
ID AA919154 standard; protein; 731 AA.  
XX  
AC AA919154;

XX 19-JUL-2000 (first entry)  
XX Human cytoskeleton associated protein 9 (CYSKP-9).  
DE Cytoskeleton associated protein; CYSKP-9; cancer; proliferative;  
KW autoimmunity; inflammatory; vesicle trafficking; neurological;  
KW cardiovascular; cell motility; reproductive; muscle disorder.  
XX Homo sapiens.  
OS  
XX Key Location/Qualifiers  
FH Modified-site 15  
FT /note= "potential phosphorylation site"  
FT Domain 40..64  
FT /note= "CAP-Gly domain"  
FT Modified-site 117  
FT /note= "potential phosphorylation site"  
FT Modified-site 125  
FT /note= "potential glycosylation site"  
FT Modified-site 131  
FT /note= "potential phosphorylation site"  
FT Modified-site 134  
FT /note= "potential glycosylation site"  
FT Modified-site 136  
FT /note= "potential phosphorylation site"  
FT Modified-site 162  
FT /note= "potential phosphorylation site"  
FT Modified-site 168  
FT /note= "potential phosphorylation site"  
FT Modified-site 205  
FT /note= "potential glycosylation site"  
FT Modified-site 219  
FT /note= "potential phosphorylation site"  
FT Modified-site 249  
FT /note= "potential phosphorylation site"  
FT Modified-site 292  
FT /note= "potential phosphorylation site"  
FT Modified-site 313  
FT /note= "potential phosphorylation site"  
FT Modified-site 390  
FT /note= "potential phosphorylation site"  
FT Modified-site 407  
FT /note= "potential phosphorylation site"  
FT Modified-site 451  
FT /note= "potential phosphorylation site"  
FT Modified-site 490  
FT /note= "potential phosphorylation site"  
FT Modified-site 551  
FT /note= "potential glycosylation site"  
FT Modified-site 559  
FT /note= "potential phosphorylation site"  
FT Modified-site 665  
FT /note= "potential phosphorylation site"  
FT Modified-site 694  
FT /note= "potential phosphorylation site"  
FT Modified-site 703  
FT /note= "potential phosphorylation site"  
XX WO200017355-A2.  
XX 30-MAR-2000.  
XX 17-SEP-1999; 99WO-US021565.  
XX 18-SEP-1998; 98US-0172226P.  
XX 27-APR-1999; 99US-0131321P.  
XX (INCY-) INCYTE PHARM INC.  
XX Lal P, Tang YT, Yue H, Hillman JL, Bandman O, Corley NC;  
XX Guegler KJ, Patterson C, Azimzai Y, Baughn MR;

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 18, 2005, 09:14:43 ; Search time 73 Seconds  
(without alignments)  
5027.885 Million cell updates/sec

Title: US-09-671-687A-3

Perfect score: 949

Sequence: 1 MSSGLMSQEKVTSYWEERI.....RLLCDAYMCVQSPMTSLYK 949

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 2105692 seqs, 386760381 residues

Word size : 0

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

A\_Geneseq\_16Dec04:\*

- 1: Geneseqp1980s:\*
- 2: Geneseqp1990s:\*
- 3: Geneseqp2000s:\*
- 4: Geneseqp2001s:\*
- 5: Geneseqp2002s:\*
- 6: Geneseqp2003as:\*
- 7: Geneseqp2003bs:\*
- 8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	606	63.9	685	4	AAM39254 Human pol
2	606	63.9	731	3	AAY91954 Human cyt
3	606	63.9	739	4	AU23747 Novel hum
4	606	63.9	953	7	ADC24816 Human bre
5	606	63.9	953	6	ADG95918 T cell ac
6	606	63.9	956	8	ADG95918 Human CYL
7	606	63.9	956	6	ADG95918 Human CYL
8	606	63.9	956	8	ADG95918 Human NF-
9	505	53.2	731	4	ADG95918 Human pro
10	505	53.2	953	8	ADG95918 T cell ac
11	505	53.2	956	8	ADG95918 T cell ac
12	363	38.3	698	4	AAM41040 Human pol
13	330	34.8	476	4	AAM95719 Human pro
14	227	23.9	558	4	AU23211 Novel hum
15	208	21.9	261	5	ABB89233 Human pol
16	148	15.6	238	5	ABB89234 Human pol
17	113	11.9	113	4	AAM14965 Peptide #
18	113	11.9	113	4	ABB33941 Peptide #
19	113	11.9	113	4	ABB33941 Peptide #
20	113	11.9	113	4	ABB27399 Peptide #
21	113	11.9	113	4	ABB27399 Peptide #
22	113	11.9	113	4	ABB19377 Protein #
23	113	11.9	113	4	AAM67104 Human bon
24	113	11.9	113	4	AAM54704 Human bra
25	113	11.9	113	4	AAM02691 Peptide #
					ABG36764 Human pep

26	106	11.2	106	8	ADK71947	Adk71947 Human ori
27	101	10.6	101	8	ADK71945	Adk71945 Human ori
28	96	10.1	96	8	ADK71949	Adk71949 Human ori
29	91	9.6	91	8	ADK71943	Adk71943 Human ori
30	91	9.6	104	8	ADK71962	Adk71962 Human ori
31	10	1.1	517	4	ABB61669	Abb61669 Drosophil
32	8	0.8	78	5	ABP68782	Abp68782 Marine sn
33	8	0.8	120	5	ADK35501	Adk35501 Novel hum
34	8	0.8	216	4	AAO72831	AAO72831 Human olf
35	8	0.8	228	7	ABO73471	ABO73471 Pseudomon
36	8	0.8	418	4	ABB64546	Abb64546 Drosophil
37	8	0.8	450	5	ABB50103	Abb50103 Listeria
38	8	0.8	450	5	ABU32881	Abu32881 Protein e
39	8	0.8	551	6	ABR83636	ABR83636 Saccharom
40	8	0.8	644	7	ADB80358	ADB80358 Human MDD
41	8	0.8	708	7	ABO71275	ABO71275 Pseudomon
42	8	0.8	941	7	ADF76391	Adf76391 Novel hum
43	8	0.8	1004	7	ADB80336	ADB80336 Human MDD
44	8	0.8	1086	8	ADN23310	Adn23310 Bacterial
45	7	0.7	18	5	AAOI5840	AAOI5840 CtxB sign

#### ALIGNMENTS

#### RESULT 1

AAM39254  
ID AAM39254 standard; protein; 685 AA.

XX AC AAM39254;

DT 22-OCT-2001 (first entry)

DE Human polypeptide SEQ ID NO 2399.

KW Human; neotropic; immunosuppressant; cytostatic; gene therapy; cancer;  
KW peripheral nervous system; neuropathy; central nervous system; CNS;  
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;  
KW leukaemia.

OS Homo sapiens.

XX WO200153312-A1.

PD 26-JUL-2001.

XX 26-DEC-2000; 2000WO-US034263.

XX 23-DEC-1999; 99US-00471275.

XX 21-JAN-2000; 2000US-00488725.

XX 25-APR-2000; 2000US-00552317.

XX 20-JUN-2000; 2000US-00598042.

XX 19-JUL-2000; 2000US-00620312.

XX 03-AUG-2000; 2000US-00653450.

XX 14-SEP-2000; 2000US-00682191.

XX 19-OCT-2000; 2000US-00693036.

XX 29-NOV-2000; 2000US-00727344.

(HYSE-) HYSEQ INC.

PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;

PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QH;

PI Zhou P, Goodrich R, Drmanac RT;

PI WPI: 2001-442253/47.

PI N-PSDB; AAI58410.

*late*

Novel nucleic acids and polypeptides, useful for treating disorders such as central nervous system injuries.

Example 4; SEQ ID NO 2399; 10078pp; English.

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: April 17, 2005, 06:05:31 ; Search time 6188 Seconds  
(without alignments)

5837.588 Million cell updates/sec

Title: US-09-671-687A-3

Perfect score: 949

Sequence: 1 MSSGLSQSKVTSPTWBERI.....RLLCDAYCMYQSPMTSLYK 949

Scoring table:

OLIGO  
Xgapop 60.0 , Xgapext 60.0  
Ygapop 60.0 , Ygapext 60.0  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 34239544 seqs, 19032134700 residues

Word size: 1

Total number of hits satisfying chosen parameters: 68477535

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp  
-Q=/cgn2.1/USPTO\_spool\_p/US09671687/runat\_15042005\_163038.11962/app\_query.fasta\_1.1095  
-DB=EST -QFMT=fastap -SUFFIX=oligo.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc  
-NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09671687@cgn 1\_15180 @runat\_15042005\_163038.11962 -NCFU=6 -ICPU=3  
-NO MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

EST:\*  
1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_hic:\*  
4: gb\_est3:\*  
5: gb\_est4:\*  
6: gb\_est5:\*  
7: gb\_est6:\*  
8: gb\_gss1:\*  
9: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	387	40.8	3496	3 AF161542	AF161542 Homo sapi
2	313	33.0	2862	9 AY406374	AY406374 Homo sapi
3	312	32.9	2862	9 AY406375	AY406375 Pan trogl
4	284	29.9	1039	4 BM457960	BM457960 AGENCOURT
5	243	25.6	1063	5 BQ433523	BQ433523 AGENCOURT
6	237	25.0	726	7 CR772310	CR772310 DKFZp468L
c 7	218	23.0	677	6 CD626856	CD626856 56076837J
8	208	21.9	697	4 BM724143	BM724143 UI-E-EOI-
9	207	21.8	624	7 CN335014	CN335014 170005319

10	200	21.1	601	7	CR789083	CR789083 DKFZp468L
11	197	20.8	596	6	CB286853	CB286853 CMD47_E07
12	189	19.9	575	5	BP226285	BP226285 BP226285
c 13	189	19.9	581	7	CF528626	CF528626 UI-1-BC1-
14	185	19.5	736	6	CD641429	CD641429 AGENCOURT
15	179	18.9	919	4	BG180100	BG180100 602329675
16	178	18.8	583	5	BP219825	BP219825 BP219825
17	175	18.4	689	6	CD639981	CD639981 AGENCOURT
18	174	18.3	783	4	BM016881	BM016881 603643383
c 19	173	18.2	644	1	AA044899	AA044899 zk75d10.8
20	170	17.9	510	2	BF724264	BF724264 bx02g11.y
c 21	155	16.3	473	1	AA833748	AA833748 od61b06.s
22	155	16.3	567	5	BP221754	BP221754 BP221754
c 23	155	16.3	767	6	CD370161	CD370161 UI-H-FT1-
24	154	16.2	573	5	BP221744	BP221744 BP221744
25	153	16.1	907	5	BQ438227	BQ438227 AGENCOURT
c 26	141	14.9	425	1	AI130924	AI130924 Qc11a01.x
c 27	139	14.6	447	1	AA152263	AA152263 Z103a04.s
28	139	14.6	473	1	AA256728	AA256728 zs22h05.r
29	137	14.4	708	6	CB466941	CB466941 732655 MA
30	135	14.2	753	4	BI910868	BI910868 603067618
31	134	14.1	628	1	AA044687	AA044687 zk75d10.x
32	134	14.1	1007	5	BU570801	BU570801 AGENCOURT
c 33	133	14.0	406	1	AA761514	AA761514 nz27d01.s
34	131	13.8	541	4	BG574981	BG574981 602598070
c 35	131	13.8	569	1	AA056485	AA056485 zk80b08.s
36	131	13.8	635	4	BM820463	BM820463 K-EST0089
37	129	13.6	433	6	CD626855	CD626855 56076837H
38	129	13.6	565	4	BM832976	BM832976 K-EST0107
39	129	13.6	654	6	CB555752	CB555752 MMSF0069
40	129	13.6	1134	5	BM904688	BM904688 AGENCOURT
41	127	13.4	873	4	BI907475	BI907475 603063577
42	125	13.2	633	4	BG756164	BG756164 602713465
43	124	13.1	582	5	BP360904	BP360904 BP360904
44	124	13.1	837	1	AUI37267	AUI37267 AUI37267
45	123	13.0	758	1	AUI22742	AUI22742 AUI22742

#### ALIGNMENTS

RESULT 1	AF161542	3496 bp	linear	HTC 22-MAY-2001
LOCUS	AF161542	3496 bp	mRNA	
DEFINITION	Homo sapiens HSPC057 mRNA, complete cds.			
ACCESSION	AF161542			
VERSION	AF161542.1	GI:6841351		
KEYWORDS	HTC.			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
AUTHORS	Zhang, Q.H., Ye, M., Wu, X.Y., Ren, S.X., Zhao, M., Zhao, C.J., Fu, G., Shen, Y., Fan, H.Y., Lu, G., Zhong, M., Xu, X.R., Han, Z.G., Zhang, J.W., Tao, J., Huang, Q.H., Zhou, J., Hu, G.X., Gu, J., Chen, S.J. and Chen, Z.			
TITLE	Cloning and functional analysis of cDNAs with open reading frames for 300 previously undefined genes expressed in CD34+ hematopoietic stem/progenitor cells			
JOURNAL	Genome Res. 10 (10), 1546-1560 (2000)			
MEDLINE	20499367			
PUBMED	11042152			
REFERENCE	2 (bases 1 to 3496)			
AUTHORS	Zhang, Q.H., Ye, M., Zhou, J., Shen, Y., Wu, X.Y., Guan, Z.Q., Wang, L., Fan, H.Y., Mao, Y.F., Dai, M., Huang, Q.H., Chen, S.J. and Chen, Z.			
TITLE	Human full length cDNA cloned from cd34+ stem cells			
JOURNAL	Unpublished			
REFERENCE	3 (bases 1 to 3496)			
AUTHORS	Zhang, Q.H., Ye, M., Zhou, J., Shen, Y., Wu, X.Y., Guan, Z.Q., Wang, L., Fan, H.Y., Mao, Y.F., Dai, M., Huang, Q.H., Chen, S.J. and Chen, Z.			
TITLE	Direct Submission			
JOURNAL	Submitted (14-MAY-1999) Shanghai Institute of Hematology, Shanghai Second Medical University, Rui-Jin Hospital, 197 Rui-Jin Road II, Shanghai 200025, P. R. China			

FEATURES  
Source

Location/Qualifiers

```
1. .3496
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CBPA0803"
/tissue_type="blood"
1517. .3064
/codon_start=1
/product="HSPC057"
/protein_id="AAP29029.1"
/db_xref="GI:5841352"
/translation="MPNTNGSIHSPSLSAQSVMEELNTPVQESPPLAMPNGSHG
LEVGSLAEVKEENPFYGVIRWIGQPPGLNEVLAGELEDECAGCTDGTFRGTRYPTCA
LKALFVYKSCRSDFSPASIQPVSNQIERCNSLAFGGYLSVVVEENTPPKMEKGLLE
IMKKGKGIQGHVNSYLDLSTFCLFAPSSVLDTVLLRPKEKNDVEYYSYEQELLRLTE
IVNPLRIVGYCANKIKMLKILKEVAAAGSTSEKDPSEFLNLFLPHILRLVPELLK
IRSAQKQDCCFYQIEMKNEKGVPTIQQLLEWSFINSLNKAFAEASCLLIQMFRF
GKDPKPKKIPPSLELIYQIYLKLPDSAGVVEGLQCMVENATTIRFSAGKIQKFC
KTCNTQVHLHPKRLNKHYNVPSLPLKDPDMDRHCIPCONNELFAVLICITSHYVAF
VKYKGDSDSAMLFFDSMADRDGQGNFIPQVTPCPVEGYLKMSLDLHSLDSRRIPR
LCTKTAL"
```

## ORIGIN

## Alignment Scores:

```
Pred. No.: 0 Length: 3496
Score: 387.00 Matches: 558
Percent Similarity: 99.11% Conservative: 0
Best Local Similarity: 99.11% Mismatches: 2
Query Match: 40.78% Indels: 5
DB: 3 Gaps: 0
```

US-09-671-687A-3 (1-949) x AF161542 (1-3496)

```
Qy 390 ArgSerSerProProLeuGlnProProValAsnSerLeuThrThrGluAsnArgPhe 409
Db 1430 CGGTCTTACACACCTCCAGCCTCTCTGTGAACCTCACTGACACCGAGAACAGATTC 1489
Qy 410 HisSerLeuProPheSerLeuThrLysMetProAsnThrAsnGlySerIleGlyHisSer 429
Db 1490 CACTCTTTTACCATTGAGTCTCACCAGATGCCAATACCAATGGAAGTATTGGCCACAGT 1549
Qy 430 ProLeuSerLeuSerAlaGlnSerValMetGluGluLeuAsnThrAlaProValGlnGlu 449
Db 1550 CCACCTTCTCTGTGAGCCAGCTCTGTANTGAGAGAGCTAAACACTGCACCCGTCAGAG 1609
Qy 450 SerProProLeuAlaMetProProGlyAsnSerHisGlyLeuGluValGlySerLeuAla 469
Db 1610 AGTCCACCTTTGGCCATGCCCTCTGGGAACCTCACATGGTCTAGAAGTGGGCTCAITGGCT 1669
Qy 470 GluValLysGluAsnProProPheThrGlyValIleArgTrpIleGlyGlnProProGly 489
Db 1670 GAAGTTAAGGAGAACCCCTCTTCTATGGGGTAAATCCGTTGGATCGGTCAAGCAGGGA 1729
Qy 490 LeuAsnGluValLeuAlaGlyLeuGluLeuAspGluCysAlaGlyCysThrAspGly 509
Db 1730 CTGAATGAGTGCTCGCTGGACTGGAACCTGGAGATGAGTGTGAGGCTGTACGAGATGA 1789
Qy 510 ThrPheArgGlyThrArgTyPheThrCysAlaLeuLysLysAlaLeuPheValLysLeu 529
Db 1790 ACCTTCAGAGGCACTCGGTATTTCACCTGTGCCCTGGAAGAGGGCGCTGTTGTGAAACTG 1849
Qy 530 LysSerCysArgProAspSerArgPheAlaSerLeuGlnProValSerAsnGlnIleGlu 549
Db 1850 AAGAGCTGCAGCCCTGACTCTAGTGTTCATTCATTCAGCCGCTTTCGAATCAGATTGAG 1909
Qy 550 ArgCysAsnSerLeuAlaPheGlyGlyTyThrLeuSerGluValValGluGluAsnThrPro 569
Db 1910 CGCTGTAACTCTTTAGCAATTTGGAGGCTACTTAAGTGAAGTAGTAGAAGAAATAACTCCA 1969
Qy 570 ProLysMetGluLysGluGluIleMetIleGlyLysLysGlyIleGlnGly 589
Db 1970 CCAAAAATGGAAAGAGGCTTGGAGATATGATTGGGAAGAGAAAGGCATCCAGGGT 2029
```

```
Qy 590 HisTyAsnSerCysTyThrLeuAspSerThrLeuPheCysLeuPheAlaPheSerVal 609
Db 2030 CATTACAATCTTGTTACTTAGACTCAACCTTATTCTGCTATTGCTTTTAGTCTGTT 2089
Qy 610 LeuAspThrValLeuLeuArgProLysGluLysAsnAspValGluTyThrSerGluThr 629
Db 2090 CTGGACACTGTGTTACTTAGACCCCAAGAAAGACGATGTAGAATAATTATAGTGAAC 2149
Qy 630 GlnGluLeuLeuArgThrGluIleValAsnProLeuArgIleTyThrGlyTyValCysAla 649
Db 2150 CAAGAGCTACTGAGACGAGAAATTGTTAATCTCTGAGAATATATGGATATGTGTGTC 2209
Qy 650 ThrLysIleMetLysLeuArgLysIleLeuGluLysValGluAlaAlaSerGlyPheThr 669
Db 2210 ACAAAAATATGAAACTGAGGAANATACTTGAAGAGTGGAGGCTGCATCAGAAATTACC 2269
Qy 670 SerGluGluLysAspProGluGluPheLeuAsnIleLeuPheHisIleLeuArgVal 689
Db 2270 TCTGAAGAAAAAGATCCTGAGGAATCTTGAATATTCTGTTTTCATCATATTTTAAGG 2329
Qy 690 GluProLeuLeuLysIleArgSerAlaGlyGlnLysValGlnAspCysTyThrPheTyGln 709
Db 2330 GAACCTTTTGTAAAAATAAGATCAGAGGTCAAAAGGTACAAAGATTGTTTACTTCTCAA 2389
Qy 710 IlePheMetGluLysAsnGluLysValGlyValProThrIleGlnGlnLeuLeuGluTrp 729
Db 2390 ATTTTATGGAATAAAATGAGNAAGTTGGCGTTCACCAATTCAGCAGTGTGTAGNATGG 2449
Qy 730 SerPheIleAsnSerAsnLeuLysPheAlaGluAlaProSerCysLeuIleIleGlnMet 749
Db 2450 TCTTTTATCAACAGTAACCTGAAATTTGCAGAGGCCACCATCATGCTGTGATTTATTCAG 2509
Qy 750 ProArgPheGlyLysAspPheLysIlePhePheLysLysIlePheProSerLeuGluLeuAsn 769
Db 2510 CCTGATTTGGAAAAAGACTTTTAAACTATTATAAAAAATTTTTCCTCTCTGGAATT-AA 2568
Qy 770 IleThr-AspLeuLeuGluAspThrProArgGlnCysArgIleCysGlyGlyLeuAlaMe 789
Db 2569 AT-ACCAGATTTTACTTTGAAGACACTCCACAGAGTCCGCGATATGTGGAGGGCTTCCA 2627
Qy 789 tTyGluCysArgGluCysTyAspAspProAspIle-SerAlaGlyLysIleLysGlnP 809
Db 2628 GTATCAGTGTAGAGAAATGCTACGACGATCCGAGCATTCACAGTGGAAAAATCAAGCAG 2687
Qy 809 heCysLysThrCysAsnThrGlnValHisLeuHisProLysArgLeuAsnHisLysTyA 829
Db 2688 TTTGTAAAAACCTGCAACACTCAAGTCCACCTTCATCCGAAGAGGCTGAATCATATAATA 2747
Qy 829 enProValSerLeuProLysAspLeuProAspTrpAspTrpArgHisGlyCysIleProC 849
Db 2748 ACCCAGTGTCACTTCCCAAGACTTACCGACTGGGACTGGAGACACGCGCTGCATCCCTT 2807
Qy 849 ysGlnAsnMetGluLeuPheAlaValLeuCysIleGluThrSerHisTyValAlaPheV 869
Db 2808 GCCAGAAATATGGAGTTATTGCTGTTCTCTGATAGAAACAAGCCACATATGTTGCTTTG 2867
Qy 869 alLysTyGlyLysAspSerAlaTrpLeuPhePheAspSerMetAlaAspArgAspG 889
Db 2868 TGAAGTATGGGAAGGACGATTCGCTGCTCTCTTTGACAGCATGGCCGATCGGAGATG 2927
Qy 889 lyGlyGlnAsnGlyPheAsnIleProGlnValThrProCysProGluValGlyGluTyTr 909
Db 2928 GTGGTTCAGAAATGGCTTCAACATTCCTCAAGTCACCCCATGCCCCAGAGATCCCAAG 2987
Qy 909 euLysMetSerLeuGluAspLeuHisSerLeuAspSerArgArgIle-GlnGlyCysAla 928
Db 2988 TGAAGATGCTCTGGAAGACCTGCAATTCCTTGGATCCAGGAGAAATCCCAAGGCTGTGA 3047
Qy 929 ArgArgLeuLeuCysAspAlaTyThrMetCysMetTyThrGlnSerProThrMetSerLeu 948
Db 3048 CGAAGACTGCTTTGTGATGCATATATGTGCATGTACAGAGTCCCAACAAATGATGTGTAC 3107
```

```

QY      949 Lys 949
Db      3108 AAA 3110

RESULT 2
LOCUS   AY406374      2862 bp      DNA      linear      GSS 15-DEC-2003
DEFINITION Homo sapiens CYLD gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION AY406374
VERSION   AY406374.1 GI:39762348
KEYWORDS GSS.
SOURCE   Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2862)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE    Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED   14671302
REFERENCE 2 (bases 1 to 2862)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE    Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT  This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source    1..2862
           /organism="Homo sapiens"
           /mol_type="genomic DNA"
           /db_xref="taxon:9606"
           <1..>2862
           /gene="CYLD"
           /locus_tag="HCM2536"
gene
gene
ORIGIN
Alignment Scores:
Pred. No.:      0      Length:      2862
Score:          313.00  Matches:      340
Percent Similarity: 99.13%  Conservative: 0
Best Local Similarity: 99.13%  Mismatches: 0
Query Match:      32.98%  Indels:      3
DB:              9      Gaps:        0

US-09-671-687A-3 (1-949) x AY406374 (1-2862)

QY      303 SerValThrGlnGluArgArgProProLysLeuAlaPheMetSerArgGlyValGlyAsp 322
Db      916 AGTGTCAGCGAGAAAGAGGCGCTCCCAACTGCTTTATGTCAGAGGCTGTGGGAC 975

QY      323 LysGlySerSerSerHisAsnLysProLysAlaThrGlySerThrSerAspProGlyAsn 342
Db      976 AAAGGTCATCCAGTCATATAAACCAAGAGGTCACAGATCTACCTCAGACCCCTGGAAT 1035

QY      343 Arg---ArgSerGluLeuPheTyrThrLeuAsnGlySerSerValAspSerGlnProGln 361
Db      1036 AGAAACAGATCTGAATATTTTATACCTTAATGGTCTCTCTGTTGACTCACACACAA 1095

QY      362 SerLysSerLysAsnThrTyrTyrIleAspGluValAlaGluAspProAlaLysSerIleu 381
Db      1096 TCCAAATCAAAAAATACATGTTACATGATGAAGTTGCAGAGAGCCCTGCAAAATCTCTT 1155

QY      382 ThrGluIleSerThrAspPheAspArgSerSerProProLysGlnProProValAsn 401

```

```

Db      1156 ACAGAGATATCTACAGACTTTGACCGCTTCTTCACCACCACCTCCAGCGCTCCTCTGTGAAC 1215
QY      402 SerLeuThrThrGluAsnArgPheHisSerLeuProPheSerLeuThrLysMetProAsn 421
Db      1216 TCACTGACCACCGAGAACAGATTCCACTCTTTTACCATTAGTCTCAACCAAGATGCCCAAT 1275
QY      422 ThrAsnGlySerIleGlyHisSerProLysSerLeuSerAlaGlnSerValMetGluGlu 441
Db      1276 ACCAATGGAAGATTATGGCCACAGTCCACTTCTGTGACCCAGCTCTGTAATGGAAGAG 1335
QY      442 LeuAsnThrAlaProValGlnGluSerProProLeuAlaMetProProGlyVAsnSerHis 461
Db      1336 CTAAACACCTGCACCCGCTCAAGAGAGTCCACCTTGGCCATGGCTCTCTGGAACTCACAT 1395
QY      462 GlyLeuGluValGlySerLeuAlaGluValLysGluAsnProProPheTyrGlyValIle 481
Db      1396 GGCTAGAGAGTGGGCTCATTTGGCTGAAGTTAAGGAGAACCCCTCCTTCTATGGGTAATC 1455
QY      482 ArgTTrpIleGlyGlnProProGlyLeuAsnGluValLeuAlaGlyLeuGluLeuGluAsp 501
Db      1456 CGTTGGATCGGTGAGCCACCGAGACTGAATGAAGTGTCTGCTGGACTGGAACTGGAAGAT 1515
QY      502 GluCysAlaGlyCysThrAspGlyThrPheArgGlyThrArgTyrPheThrCysAlaLeu 521
Db      1516 GAGTGTGACGGCTGTACGGATGGAACCTTCAGAGGCACCTCGGTATTTTCACTGTGCCCTG 1575
QY      522 LysLysAlaLeuPheValLysLeuLysSerCysArgProAspSerArgPheAlaSerLeu 541
Db      1576 AAGAGGCGCTGTTTGTGAACCTGAAGAGCTGAGGCGTGAAGCTCTAGTTTGCATCATTTG 1635
QY      542 GlnProValSerAsnGlnIleGluArgCysAsnSerLeuAlaPheGlyGlyTyrLeuSer 561
Db      1636 CAGCGGTTTCCAATCAGATTGAGCGCTGTAACTCTTTAGCACTTGGAGCTACTTAAAT 1695
QY      562 GluValValGluGluAsnThrProProLysMetGluLysGluGlyLeuGluIleMetIle 581
Db      1696 GAAGTAGTAGAAGAAATATCTCCACCAAAATGSAAGAGAGCGCTGGAGATAATGATT 1755
QY      582 GlyLysLysGlyIleGlnGlyHisTyrAsnSerCysTyrLeuAspSerThrLeuPhe 601
Db      1756 GGGAAGAAAGAAAGGCAATCCAGGTCATTACAAATCTTGTGTACTTAGACTCAACTTATTC 1815
QY      602 CysLeuPheAlaPheSerValLeuAspThrValLeuLeuArgProLysGluLysAsn 621
Db      1816 TGTCTATTGCTTTTAGTTCTGTTCTGGACACTGTGTACTTAGACCAAGAAAGAAAGAAC 1875
QY      622 AspValGluTyrTyrSerGluThrGlnGluLeuLeuArgThrGluIleValAsnProLeu 641
Db      1876 GATGTAGAATATTATAGTGAACCCCAAGAGCTACTGAGGACAGAAATTTGTTAATCTCTG 1935
QY      642 Arg 642
Db      1936 AGA 1938

RESULT 3
AY406375      2862 bp      DNA      linear      GSS 15-DEC-2003
LOCUS   AY406375
DEFINITION Pan troglodytes CYLD gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION AY406375
VERSION   AY406375.1 GI:39762349
KEYWORDS GSS.
SOURCE   Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE 1 (bases 1 to 2862)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE    Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios

```

JOURNAL Science 302 (5652), 1960-1963 (2003)  
PUBMED 14671302  
REFERENCE 2 (bases 1 to 2862)  
AUTHORS Clark,A.G., Ghanowski,S., Nielson,R., Thomas,P., Kejarival,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriter,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.O., Adams,M.D. and Cargill,M.  
TITLE Direct Submission  
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA  
COMMENT This sequence was made by sequencing genomic exons and ordering them based on alignment.  
FEATURES  
source Location/Qualifiers  
1..2862  
/organism="Pan troglodytes"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9598"  
<1..>2862  
/gene="CYLD"  
/locus\_tag="HCW2536"  
ORIGIN  
Alignment Scores:  
Pred. No.: 0 Length: 2862  
Score: 312.00 Matches: 339  
Percent Similarity: 99.12% Conservative: 0  
Best Local Similarity: 99.12% Mismatches: 0  
Query Match: 32.88% Indels: 3  
DB: 9 Gaps: 0  
US-09-671-687A-3 (1-949) x AY406375 (1-2862)  
QY 304 ValThrGlnGluArgArgProProlysisLeuAlaPheMetSerArgGlyValGlyAspLys 323  
DB 919 GTGACGCGAGGAAGAGGCGCTCCCAACTTGTCTTATGTCAAGAGGTGTGGGACAAA 978  
QY 324 GlySerSerHisAsnLysProLysAlaThrGlySerThrSerAspProGlyAsnArg 343  
DB 979 GGTTCATCCAGTCATAATAAACCAAGGCTACAGGATCTACCTCAGACCTCGGAATAGA 1038  
QY 344 ---ArgSerGluLeuPheTyrThrLeuAsnGlySerSerValAspSerGlnProGlnSer 362  
DB 1039 AACAGATCTGAATATATTTATACCTTAATGGGTCTTCTGTGTGACTCAACCAATCC 1098  
QY 363 LysSerLysAsnThrTrpTyrIleAspGluValAlaGluAspProAlaLysSerLeuThr 382  
DB 1099 AAATCAAAAATAACATACATGATGAAGTTGCAGAGACCCCTGCAGAAATCTCTTACA 1158  
QY 383 GluIleSerThrAspPheAspArgSerProProLeuGlnProProValAsnSer 402  
DB 1159 GAGATATCTACAGACTTTGACCGTTCTTCACCAACCACTCCAGCTCTCTGTGAACCTA 1218  
QY 403 LeuThrThrGluAsnArgPheHisSerLeuProPheSerLeuThrLysMetProAsnThr 422  
DB 1219 CTGACCCGAGAACAGATTCACCTCTTTACCATTCAGTCTCACCAGATGCCAATACC 1278  
QY 423 AsnGlySerIleGlyHisSerProLeuSerLeuSerAlaGlnSerValMetGluGluLeu 442  
DB 1279 AATGGAAGTATTTGGCCACAGTCCACTTTCTGTGAGCCAGTCTGTATGGAAGAGCTA 1338  
QY 443 AsnThrAlaProValGlnGluSerProProLeuAlaMetProProGlyAsnSerHisGly 462  
DB 1339 AACACTGACCCGTCACAGAGAGTCCACCTTTGGCCATGCTCTCTGGGAACCTCAGATGGT 1398  
QY 463 LeuGluValGlySerLeuAlaGluValLysGluAsnProProPheTyrGlyValIleArg 482  
DB 1399 CTAGAAGTGGGCTCATTTGGCTGAAGTTAAGGAGAACCTCTCTTCTATGGAGTATCCGT 1458  
QY 483 TrpIleGlyGlnProProGlyLeuAsnGluValLeuAlaGlyLeuGluLeuGluAspGlu 502  
DB 1459 TGGATCGGTGACCCACAGGAGTGAATGAAGTCTCGCTGGACTGGAACTGGAGATGAG 1518  
QY 503 CysAlaGlyCysThrAspGlyThrPheArgGlyThrArgTyrPheThrCysAlaLeuLys 522

Db 1519 TGTGAGGCTGTACGATGGAACCTTTCAGAGGCACTCGGTATTTTCCCTGTGCTGAAG 1578  
QY 523 LysAlaLeuPheValLysLeuLysSerCysArgProAspSerArgPheAlaSerLeuGln 542  
Db 1579 AAGGCGCTGTTTGTGAAACTGAAGAGCTGCAGGCTGACTCTAGGTTTTCATTCAGAG 1638  
QY 543 ProValSerAsnGlnIleGluArgCysAsnSerLeuAlaPheGlyGlyTyrLeuSerGlu 562  
Db 1639 CGGTTTCCATCAGATCGAGCGCTGTAACCTTTAGCATTTGGAGGCTACTTAAGTGA 1698  
QY 563 ValValGluGluAsnThrProProLysMetGluLysGluGlyLeuGluLeuMetIleGly 582  
Db 1699 GTAGTAGAAGAAATACTCCACCAAAATGGAAGAAAGAGGCTTGAGAGATAATGATTGG 1758  
QY 583 LysLysLysGlyIleGlnGlyHisTyrAsnSerCysTyrLeuAspSerThrLeuPheCys 602  
Db 1759 AAGAAGAAGGATCCAGGCTCATTAACATTTCTGTACTTAGACTCAACCTTATTCTGC 1818  
QY 603 LeuPheAlaPheSerSerValLeuAspThrValLeuLeuArgProLysGluLysAsnArg 622  
Db 1819 TTATTTGCTTTTGTCTGTCTGACACTGTGTACTTAGACCCCAAGAAAGACGAT 1878  
QY 623 ValGluTyrTyrSerGluThrGlnGluLeuLeuArgThrGluIleValAsnProLeuArg 642  
Db 1879 GTAGAATATTATAGTGAACCAAGAGCTACTGAGACAGAAATGTTAATCTCTGAGA 1938  
RESULT 4  
BM457960 1039 bp mRNA linear EST 05-FEB-2002  
LOCUS AGENCOURT 6411653 NIH\_MGC\_71 Homo sapiens cDNA clone IMAGE:5530295  
DEFINITION 5', mRNA sequence.  
ACCESSION BM457960  
VERSION BM457960.1 GI:18507000  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 1039)  
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs@mail.nih.gov  
Tissue Procurement: ATCC  
cDNA Library Preparation: Life Technologies, Inc.  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLAM12210 row: b column: 24  
High quality sequence stop: 702.  
FEATURES  
Location/Qualifiers  
1..1039  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5530295"  
/tissue\_type="leiomyosarcoma"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH MGC\_71"  
/note="Organ: uterus; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dr.  
Average insert size 2.1 kb."  
ORIGIN  
Alignment Scores:  
Pred. No.: 3,356-299 Length: 1039  
Score: 284.00 Matches: 284  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0

```
Query Match: 29.93% Indels: 0
DB: 4 Gaps: 0
US-09-671-687A-3 (1-949) x BM457960 (1-1039)

Qy 547 GlnIleGluArgCysAsnSerLeuAlaPheGlyTyrLeuSerGluValValGluGlu 566
Db 14 CAGATTGAGCGCTGTAACCTTTAGCATTTGGAGGCTACTTAAGTGAAGTAGTAGAAGAA 73

Qy 567 AsnThrProProLysMetGluLysGluGlyLeuGluIleMetIleGlyLysLysGly 586
Db 74 AATACTCCACCAAAATGCAAAAGAGGCTTGGAGATAATGATTGGGAAGAGAAAGGC 133

Qy 587 IleGlnGlyHisTyrAsnSerCysTyrLeuAspSerThrLeuPheCysLeuPheAlaPhe 606
Db 134 ATCCAGGGCTAATACAATCTTGTACTTAGACTCAACCTTATTCGCTATTGCTTTT 193

Qy 607 SerSerValLeuAspThrValLeuLeuArgProLysGluLysAsnAspValGluTyrTyr 626
Db 194 AGTTCTGTTCTGGACACTGTGTTACTTAGACCCCAAGAAAGAACGATGTAGATATTAT 253

Qy 627 SerGluThrGlnGluLeuLeuArgThrGluIleValAsnProLeuArgIleTyrGlyTyr 646
Db 254 AGTGAACCCCAAGAGCTACTGAGGACAGAAATGTTAATCTCTCAGAAATATATGATAT 313

Qy 647 ValCysAlaThrIleMetLysLeuArgLysIleLeuGluLysValGluAlaLaser 666
Db 314 GTGTGTGCCCAAAATATGAACTGAGGAAATACCTTGAAGAGTGGAGGCTGCATCA 373

Qy 667 GlyPheThrSerGluGluLysAspProGluGluPheLeuAsnIleLeuPheHisIle 686
Db 374 GGATTACTCTGAAGAAAGAGATCTGAGGAATCTTGAATATCTGTTTCATCATATT 433

Qy 687 LeuArgValGluProLeuLeuLysIleArgSerAlaGlyGlnLysValGlnAspCysTyr 706
Db 434 TTAAGGGTAGAACCTTGTCTAAAAATAAGATCAGCAGGTCAAAAGGTACAAGATTGTAC 493

Qy 707 PheTyrGlnIlePheMetGluLysAsnGluLysValGlyValProThrIleGlnGlnLeu 726
Db 494 TTCATCAATTTTATGGAATAAATGAGAAAGTTGGCGTTCCCACAATTCAGCAGTTG 553

Qy 727 LeuGluTTPSerPheIleAsnSerAsnLeuLysPheAlaGluAlaProSerCysLeuIle 746
Db 554 TTAGATGGCTTTTATCAACAGTAACCTGAAATTTGCAGAGGCACCATCTCGATT 613

Qy 747 IleGlnMetProArgPheGlyLysAspPheLysLeuPheLysLysIlePheProSerLeu 766
Db 614 ATTCAGATGCCTCGATTTGGAAAGACATTTAAACTATTATAAAAAATTTTCTCTCTG 673

Qy 767 GluLeuAsnIleThrAspLeuLeuGluAspThrProArgGlnCysArgIleCysGlyGly 786
Db 674 GAATTAATAATAACAGATTACTTGAAGACACTCCAGACAGTCGCGGATATGTGAGGG 733

Qy 787 LeuAlaMetTyrGluCysArgGluCysTyrAspAspProAspIleSerAlaGlyLysIle 806
Db 734 CTTCAATATGATAGGTAGAGATGCTACGACGATCCGGACATCTCAGCTGGAATAATC 793

Qy 807 LysGlnPheCysLysThrCysAsnThrGlnValHisLeuHisProLysArgLeuAsnHis 826
Db 794 AAGCAGITTTGTAAACCTGCAACACTCAAGTCCACCTTCATCCGAGAGGCTGAATCAT 853

Qy 827 LysTyrAsnPro 830
Db 854 AAATATAACCC 865

RESULT 5
BQ433523
LOCUS DEFINITION BQ433523 1063 bp mRNA linear EST 24-MAY-2002
5', mRNA sequence.
ACCESSION BQ433523
VERSION BQ433523.1 GI:21172599
KEYWORDS EST.
```

```
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1063)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM13211 row: 1 column: 06
High quality sequence stop: 672.
FEATURES
Location/Qualifiers
source 1..1063
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6015509"
/tissue_type="embryonal carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_92"
/note="Organ: testis; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 2.5 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."
```

## ORIGIN

```
Alignment Scores: 2,69e-254 Length: 1063
Pred. No.: 243.00 Matches: 243
Score: 243.00 Conservativeness: 0
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 100.00% Indels: 0
Query Match: 25.61% Gaps: 0
DB: 5

US-09-671-687A-3 (1-949) x BQ433523 (1-1063)

Qy 510 ThrPheArgGlyThrArgTyrPheThrCysAlaLeuLysLysAlaLeuPheValLysLeu 529
Db 1 ACCTTCAGAGGCACTCGGTATTTTCACCTGTGCTGCTGAAGAAGCGCTGTTGTGAAACTG 60

Qy 530 LysSerCysArgProAspSerArgPheAlaSerLeuGlnProValSerAsnGlnIleGlu 549
Db 61 AAGAGCTGCGAGGCTGACTCTAGGTTTGCATCATTCGAGCCGGTTTCCAATCAGATTGAG 120

Qy 550 ArgCysAsnSerLeuAlaPheGlyGlyTyrLeuSerGluValValGluGluAsnThrPro 569
Db 121 CGCTGTAACCTCTTACATTTGGAGGCTACTTAAGTGAAGTAGTAGAAGAAATACCTCCA 180

Qy 570 ProLysMetGluLysGluGlyLeuGluIleMetIleGlyLysLysGlyIleGlnGly 589
Db 181 CCAAAATGGAAAAAGAGGCTTGGAGATAATGATTGGGAAGAAAGAGCATCCAGGGT 240

Qy 590 HisTyrAsnSerCysTyrLeuAspSerThrLeuPheCysLeuPheAlaPheSerSerVal 609
Db 241 CATTACAATTTCTTACTTAGACTCAACTTATTCGCTTATTTGCTTTTAGTCTGTT 300

Qy 610 LeuAspThrValLeuLeuArgProLysGluLysAsnAspValGluTyrTyrSerGluThr 629
Db 301 CTGGACACTGTGTTACTTAGACCCCAAGAAAGAGCATGTAGAAATATTATAGTGAACCC 360

Qy 630 GlnGluLeuLeuArgThrGluIleValAsnProLeuArgIleTyrGlyTyrValCysAla 649
Db 361 CAAGAGCTACTGAGGACAGAAATTTTAATCTCTCAGAAATATATGATATGTGTGTC 420
```



```

QY 650 ThrlylMetLysLeuArgLysIleLeuGluLysValGluAlaAlaSerGlyPheThr 669
Db 421 ACAAATTTATGAACCTGAGGAAATCTTGAAAGGTTGGAGCTGCATCAGGATTTACC 480
QY 670 SerGluLysAspProGluGluPheLeuAsnIleLeuPheHisIleLeuArgVal 689
Db 481 TCTGAAGAAAAGATCTCTGAGGAATCTCTGTAATATTTCTGTTTCATCATATTTAAGGGTA 540
QY 690 GluProLeuLysIleAlaGlyGlnLysValGlnAspCysTyPheTyGln 709
Db 541 GAACCTTTGTCTAAATAAGATCAGCAGTCAAGGTCAGAGATTGTTCTCTATCA 600
QY 710 IlePheMetGluLysAsnGluLysValGlyValProThrIleGlnLeuLeuGluTrp 729
Db 601 ATTTTATGAAAAAATAGAAAAGTTGGCGTTCACAAATTCAGCAGTTGTTAGAAATGG 660
QY 730 SerPheIleAsnSerAsnLeuLysPheAlaGluAlaProSerCysLeuIleGlnMet 749
Db 661 TCTTTTATCAACAGTAACCTGAATTTTGCAGAGCACCATCATGCTCTGATTATTCAGATG 720
QY 750 ProArgPhe 752
Db 721 CTTCTGATTT 729

RESULT 6
LOCUS CR772310
DEFINITION DKFZp468L0123_r1 468 (synonym: phrt1) Pongo pygmaeus CDNA clone
ACCESSION DKFZp468L0123_5, mRNA sequence.
VERSION CR772310
KEYWORDS EST.
SOURCE Pongo pygmaeus (orangutan)
ORGANISM Pongo pygmaeus
REFERENCE 1 (bases 1 to 726)
AUTHORS Poustka,A., Albert,R., Moosmayer,P., Schupp,I., Wellenreuther,R.,
Mewes,H.W., Weill,B., Amid,C., Osanger,A., Fobo,G., Han,M. and
Wiemann,S.
TITLE Pongo pygmaeus mRNA (Poustka,A., Albert,R., Moosmayer,P., et al.)
JOURNAL Unpublished (2004)
COMMENT Contact: MIPS
MIPS Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert. Clone from S. Wiemann,
Molecular Genome Analysis, German Cancer Research Center (DKFZ);
Email s.wiemann@dkfz-heidelberg.de; any. Please contact RZPD for
ordering:
http://www.rzpd.de/cgi-bin/products/cl.cgi?CloneID=DKFZp468L0123
Further information about the clone and the sequencing project is
available at http://mips.gsf.de/projects/cdna/.
FEATURES
Location/Qualifiers
source
1..726
/organism="Pongo pygmaeus"
/mol_type="mRNA"
/db_xref="taxon:9600"
/clone="DKFZp468L0123"
/tissue_type="heart"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="468 (synonym: phrt1)"
/notes="Vector: pSport1_Sfi; Site_1: SfiI; Site_2: SfiIb"

ORIGIN
Alignment Scores:
Pred. No.: 6.77e-248 Length: 726
Score: 237.00 Matches: 237
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 24.97% Indels: 0
DB: 7 Gaps: 0

```

```

US-09-671-687A-3 (1-949) x CR772310 (1-726)
QY 525 LeuPheValLysLeuLysSerCysArgProAspSerArgPheAlaSerLeuGlnProVal 544
Db 2 CTGTTTGTAACCTGAGAGCTGCGAGGCTGACTCTAGTTTGCATCATTTGACGCGGTT 61
QY 545 SerAsnGlnIleGluArgCysAsnSerLeuAlaPheGlyGlyTyLeuSerGluValVal 564
Db 62 TCCAATCAGATTGAGCGCTGTAACCTCTTTAGCATTTGGAGGCTACTTAAGTCAAGTAGTA 121
QY 565 GluGluAsnThrProProLysMetGluLysGluGlyLeuGluIleMetIleGlyLysLys 584
Db 122 GAAGAAAATATCTCCCAAAAAATGAAAAAGAGCCCTGGAGATAATGATTGGGAAGAAG 181
QY 595 LysGlyIleGlnGlyHisTyAsnSerCysTyLeuAspSerThrLeuPheCysLeuPhe 604
Db 192 AAAGCATCCAGGGTCATTACAAATTTCTTTAGTACTCAACCTTATCTGCTTATTT 241
QY 605 AlaPheSerSerValLeuAspThrValLeuLeuArgProLysGluLysAsnAspValGlu 624
Db 242 GCTTTTAGTTCTGTTCTGACACTCTGTTACTTAGACCCAAAGAAAGACGATGAGAA 301
QY 625 TyrTySerGluThrGlnGluLeuLeuArgThrGluIleValAsnProLeuArgIleTy 644
Db 302 TATTATAGTGAACCCCAAGAGCTACTGAGGACAGAAATGTTAATCTCTGAGAAATAT 361
QY 645 GlyTyValCysAlaThrLysIleMetLysLeuArgLysIleLeuGluLysValGluAla 664
Db 362 GGATATGTTGTGTCACAAAAATATGAACTGAGGAAATAATCTTCAAAAGGTGGAGGT 421
QY 665 AlaSerGlyPheThrSerGluLysAspProGluGluPheLeuAsnIleLeuPheHis 684
Db 422 GCATCAGGATTTACCTCTGAAGAAAAGATCTCTGAGGAATCTTGAATATCTCTGTTCA 481
QY 685 HisIleLeuArgValGluProLeuLeuLysIleArgSerAlaGlyGlnLysValGlnAsp 704
Db 482 CATATTTTAAGGGTAGAACCTTTGCTAAAAATAAGATCAGCAGGTCAAAAGGTACAAGAT 541
QY 705 CysTyPheTyThrGlnIlePheMetGluLysAsnGluLysValGlyValProThrIleGln 724
Db 542 TGTTACTTCTATCAAAATTTTATGGAAGAAAAATGAGAAAGTTGGAGTTCCCAATTCAG 601
QY 725 GlnLeuLeuGluTyrSerPheIleAsnSerAsnLeuLysPheAlaGluAlaProSerCys 744
Db 602 CAGTTGTTAGATAGTGTCTTTTATCAACAGTAACCTTGAATTTTGCAGAGCACCATCAT 661
QY 745 LeuIleIleGlnMetProArgPheGlyLysAspPheLysLeuPheLysLys 761
Db 662 CTGATTATTCAGATGCTCGATTGGAAGAAAGACTTTAAACTATTATTAATAA 712

RESULT 7
LOCUS CD626856
DEFINITION 56076837J1 FLP Homo sapiens cDNA, mRNA sequence.
ACCESSION CD626856
VERSION CD626856.1 GI:40275122
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 677)
AUTHORS Fu,G.K., Wang,J.T., Yang,J., Au-Young,J. and Stuve,L.I.
TITLE Circular rapid amplification of cDNA ends for high-throughput
extension cloning of partial genes
JOURNAL Genomics 84 (1), 205-210 (2004)
COMMENT Contact: Fu GK
Incyte Genomics, Inc.
3160 Porter Dr., Palo Alto, CA 94304, USA
Tel: 6508454102
Email: gfu@incyte.com.
FEATURES
Location/Qualifiers
source
1..677

```

/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone\_lib="FLP"  
/note="Vector: pDrive Cloning Vector"

## ORIGIN

## Alignment Scores:

Pred. No.: 4,02e-227 Length: 677  
Score: 218.00 Matches: 218  
Percent Similarity: 100.00% Conservatives: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 22.97% Indels: 0  
DB: 6 Gaps: 0

US-09-671-687A-3 (1-949) x CD626856 (1-677)

Qy 619 GluLysAsnAspValGluTyrTyrSerGluThrGlnGluLeuLeuArgThrGluLeuVal 638  
Db 671 GAAAGAACGATGATGATATATATAGTGAACCCAGAGCTACTGAGGACAGAAATGTT 612  
Qy 639 AsnProLeuArgIleTyrGlyTyrValCysAlaThrLysIleMetLysLeuArgLysIle 658  
Db 611 AATCCTCTGAGAAATATGATGATGTGTGCCACAAAATATGAACTGAGGAAATA 552  
Qy 659 LeuGluLysValGluAlaSerGlyPheThrSerGluGluLysAspProGluGluPhe 678  
Db 551 CTTGAAAGGTGAGGCTGATCAGATTTTACCTCTGAAGAAAAGATCCTGAGGAATTC 492  
Qy 679 LeuAsnIleLeuPheHisIleLeuArgValGluProLeuLeuLysIleArgSerAla 698  
Db 491 TTGAATATTCGTTTCATCATATTTTAAGGGTAGAACCTTTGCTAAATAATAGATCAGCA 432  
Qy 699 GlyGlnLysValGlnAspCysTyrPheTyrGlnIlePheMetGluLysAsnGluLysVal 718  
Db 431 GGTCAAAAGGTACAAGATGTTTACTCTATCAAAATTTTATGCAAAAAAATGAGAAGTT 372  
Qy 719 GlyValProThrIleGlnLeuLeuGluTyrSerPheIleAsnSerAsnLeuLysPhe 738  
Db 371 GCGGTTCCTCCCAATTCAGAGTGTGTAGAAATGCTTTTATCAACAGTAGAACCTGAAATTT 312  
Qy 739 AlaGluAlaProSerCysLeuIleLeuGlnMetProArgPheGlyLysAspPheLysLeu 758  
Db 311 GCAGAGGACCATCATGCTGATTTATTCAGATGCTCGATTTGGAAGAACTTTAAACTA 252  
Qy 759 PheLysLysIlePheProSerLeuGluLeuAsnIleThrAspLeuLeuGluAspThrPro 778  
Db 251 TTTAAAAAAATTTTCTCTCTGGAATTTAATATAACAGATTTTACTTGAAGACACTCC 192  
Qy 779 ArgGlnCysArgIleCysGlyGlyLeuAlaMetTyrGluCysArgGluCysTyrAspAsp 798  
Db 191 AGACAGTCCCGATATGTGGAGGGCTTGCATTTATGATGATGATGATGATGATGATGAT 132  
Qy 799 ProAspIleSerAlaGlyLysIleLysGlnPheCysLysThrCysAsnThrGlnValHis 818  
Db 131 CCGACATCTCAGCTGGAAAAATCAAGCAGATTTTGTAAACCTGCAACTCAAGTCCAC 72  
Qy 819 LeuHisProLysArgLeuAsnHisLysTyrAsnProValSerLeuProLysAsp 836  
Db 71 CTTTCATCCGAGAGGCTGAATCATATATATAATACCCAGTGTCTACTCCCAAGAC 18

## RESULT 8

## BM724143

## LOCUS

DEFINITION UI-E-E01-aiy-d-03-0-UI.r1 UI-E-E01 Homo sapiens cdna clone

ACCSSION BM724143

VERSION BM724143.1 GI:19045474

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
MEDLINE  
PUBMED  
COMMENT

1 (bases 1 to 697)  
Bonaldo,M.F., Lennon,G. and Soares,M.B.  
Normalization and subtraction: two approaches to facilitate gene  
discovery  
Genome Res. 6 (9), 791-806 (1996)  
97044477  
889548  
Contact: Soares, MB  
Coordinated Laboratory for Computational Genomics  
University of Iowa  
375 Newton Road , 4156 MEBRF, Iowa City, IA 52242, USA  
Tel: 319 335 8250  
Fax: 319 335 9565  
Email: bento-soares@uiowa.edu  
Tissue Procurement: Dr. Gregg Hageman  
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: Researchers may obtain clones from Research  
Genetics (www.resgen.com).  
Seq primer: M13 Reverse

## FEATURES

## source

1..697  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="UI-E-E01-aiy-d-03-0-UI"  
/tissue\_type="fetal eye"  
/dev\_stage="fetal"  
/lab\_host="DH10B (Life Technologies) (T1 phage resistant)"  
/clone\_lib="UI-E-E01"  
/note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a  
modified polylinker; Site 1: EcoR I; Site 2: Not I;  
UI-E-E01 is a normalized cDNA library containing the  
following tissue(s): fetal eye. The library was  
constructed according to Bonaldo, Lennon and Soares,  
Genome Research, 6:791-806, 1996. First strand cDNA  
synthesis was primed with an oligo-dT primer containing a  
Not I site. Double stranded cDNA was ligated to an EcoR I  
adaptor, digested with Not I, and cloned directionally  
into pT7T3-Pac vector. The oligonucleotide used to prime  
the synthesis of first-strand cDNA contains a library tag  
sequence that is located between the Not I site and the  
(dT)18 tail. The sequence tag for this library is  
CCGATATACC. This library was created for the program, Gene  
Discovery in the Visual System, supported by National Eye  
Institute (NEI)."

## ORIGIN

## Alignment Scores:

Pred. No.: 3.69e-216 Length: 697  
Score: 208.00 Matches: 208  
Percent Similarity: 100.00% Conservatives: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 21.92% Indels: 0  
DB: 4 Gaps: 0

US-09-671-687A-3 (1-949) x BM724143 (1-697)

Qy 742 ProSerCysLeuIleIleGlnMetProArgPheGlyAspPheLysLeuPheLysLys 761  
Db 18 CCATCATGTCTGATATTAGATGCTCGATTTGGAAAAAGACTTTTAACTATTATTAATA 77  
Qy 762 IlePheProSerLeuGluLeuAsnIleThrAspLeuGluAspThrProArgGlnCys 781  
Db 78 ATTTTCTCTCTGGAATTAATATAACAGATTTTACTTGAAGACACTCCACACAGTGC 137  
Qy 782 ArgIleCysGlyGlyLeuAlaMetTyrGluCysArgGluCysTyrAspAspProAspIle 801  
Db 138 CGGATATGTGGAGGGCTTGCATATGATGATGATGATGATGATGATGATGATGATGATG 197  
Qy 802 SerAlaGlyLysIleLysGlnPheCysLysThrCysAsnThrGlnValHisLeuHisPro 821

```

Db      198 TCAGCTGAAAAAATCAACGAGTTTGTAAACCTGCAACACTCAAGTCCACCTTCATCGG 257
Qy      822 LysArgLeuAnHisLysTyAsnProValSerLeuProLysAspLeuProAspTrpAsp 841
Db      258 AAGAGGCTGAATCATAATATAACCCAGTGTACCTTCCCAAGACTTACCCGACTGGGAC 317
Qy      842 TrpArgHisGlyCysIleProCysGlnAnMetGluLeuPheAlaValLeuCysIleGlu 861
Db      318 TGGAGACACGGCTGCATCCCTTGCAGAAATATGGAGTTATTTCCTGCTCTGCATAGAA 377
Qy      862 ThrSerHisTyValAlaPheValLysTyGlyLysAspAspSerAlaTrpLeuPhePhe 881
Db      378 ACAAGCCACTATGTCTTTGTGAAGTATGGAAGGACGATCTGCGCTGGCTCTCTCTT 437
Qy      882 AspSerMetAlaAspArgAspGlyGlnAnGlyPheAsnIleProGlnValThrPro 901
Db      438 GACAGCATGGCCGATCGGATGGTGTGAGATGGCTTCACATTCCTCAGTCAACCCA 497
Qy      902 CysProGluValGlyGlyTyLeuLysMetSerLeuGluAspLeuHisSerLeuAspSer 921
Db      498 TGCCACAGAGTAGGAGAGTACTTGAAGATGCTCTGGAAGACCTGCATTCCTTGGACTCC 557
Qy      922 ArgArgIleGlnGlyCysAlaArgArgLeuLeuCysAspAlaTyMetCysMetTyArgL 941
Db      558 AGGAGATCCAAAGGCTGTGCAGAAAGACTGCTTTGTGATGCATATATGTGCATGTACCAG 617
Qy      942 SerProThrMetSerLeuTyLys 949
Db      618 AGTCCACNATGAGTTGTACAAA 641

```

## RESULT 9

```

CN335014
LOCUS      17000531936046 GRN_EB Homo sapiens cDNA 5', mRNA sequence.
DEFINITION
ACCESSION  CN335014
VERSION     CN335014.1 GI:47334948
KEYWORDS    EST.
SOURCE      Homo sapiens
ORGANISM    Homo sapiens

```

```

REFERENCE  1 (bases 1 to 624)
AUTHORS    Brandenberger, R., Wei, H., Zhang, S., Lei, S., Murage, J., Fisk, G.J.,
            Li, Y., Xu, C., Fang, R., Guegler, K., Rao, M.S., Mandalam, R.,
            Lebkowski, J. and Stanton, L.W.

```

```

TITLE      Transcriptional characterization elucidates signaling networks that
            control human ES cell growth and differentiation
JOURNAL     Nat. Biotechnol. 22 (6), 707-716 (2004)
COMMENT     Contact: Brandenberger R
            Regenerative Medicine
            Geron Corporation
            230 Constitution Drive, Menlo Park, CA 94025, USA
            Tel: 650 473 8658
            Fax: 650 473 7760
            Email: rbrandenberger@geron.com
            Insert Length: 624 Std Error: 0.00.

```

```

FEATURES   Location/Qualifiers
            source          1..624
                        /organism="Homo sapiens"
                        /mol_type="mRNA"
                        /db_xref="taxon:9606"
                        /tissue_type="embryonic stem cells, embryoid bodies
                        derived from H1, H7 and H9 cells"
                        /clone_lib="GRN_EB"
                        /note="Oligo dr primed, full-length enriched cDNA library
                        from embryoid body outgrowths derived from hES cell lines
                        H1 (p32), H7 (p29), and H9 (p26) maintained in feeder-free
                        conditions."

```

## FEATURES

```

source          1..624
                        /organism="Homo sapiens"
                        /mol_type="mRNA"
                        /db_xref="taxon:9606"
                        /tissue_type="embryonic stem cells, embryoid bodies
                        derived from H1, H7 and H9 cells"
                        /clone_lib="GRN_EB"
                        /note="Oligo dr primed, full-length enriched cDNA library
                        from embryoid body outgrowths derived from hES cell lines
                        H1 (p32), H7 (p29), and H9 (p26) maintained in feeder-free
                        conditions."

```

## ORIGIN

```

Alignment Scores:
Pred. No.:      4.11e-215      Length:      624
Score:          207.00      Matches:      207

```

```

Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match: 21.81%      Indels: 0
DB: 7      Gaps: 0

```

```

US-09-671-687A-3 (1-949) x CN335014 (1-624)

```

```

Qy      374 AlaGluAspProAlaLysSerLeuThrGluLysSerThrAspPheAspArgSerPro 393
Db      2 GCAGAAAGACCCCTGCAAAATCTCTTACAGAGATATCTACAGACTTTGACCGTTCTTCCACCA 61
Qy      394 ProLeuGlnProProProValAsnSerLeuThrThrGluAsnArgPheHisSerLeuPro 413
Db      62 CCACCTCCAGCCTCCTCTGAACTCACTACACCGAGAAACAGATTCCACTCTTTACCA 121
Qy      414 PheSerLeuThrLysMetProAsnThrAnGlySerIleGlyHisSerProLeuSerLeu 433
Db      122 TTCAGTCTCACCACAGATGCCAATACCAATGAAGTATTGGCCACAGTCCACTTCTCTG 181
Qy      434 SerAlaGlnSerValMetGluLeuAsnThrAlaProValGlnGluSerProLeu 453
Db      182 TCAGCCCACTCTGTATTAATGGAAGAGCTAAACACTGCACCCGTCCAAGAGAGTCCACCTTG 241
Qy      454 AlaMetProProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLysGlu 473
Db      242 GCATGCTCTCTGGGAACTCATCATGCTAGAAAGGGCTCATTTGGCTGAAGTAAAGGAG 301
Qy      474 AsnProProPheTyGlyValIleArgTrpIleGlyGlnProProGlyLeuAsnGluVal 493
Db      302 AACCTCTCTTCTATGGGGTAAATCCGTTGGATCGGTGAGCCACAGGACTGAATGAAGTG 361
Qy      494 LeuAlaGlyLeuGluLeuGluAspGluCysAlaGlyCysThrAspGlyThrPheArgGly 513
Db      362 CTCGCTGGACTGGAAGTGAAGTGAAGTGTGAGGCTGTACGATGGAACCTTTCAGAGGC 421
Qy      514 ThrArgTyPheThrCysAlaLeuLysLysAlaLeuPheValLysLeuLysSerCysArg 533
Db      422 ACTCGGTATTTTCACTGTGCTGCTGAAGAGGGCTGTTTGTGAAGTGAAGAGTGCAGG 481
Qy      534 ProAspSerArgPheAlaSerLeuGlnProValSerAsnGlnIleGluArgCysAsnSer 553
Db      482 CTTGACTCTAGGCTTTGCATCATTTGAGCGGGTTTCCAAATCAGATTGAGCGCTGTAACCT 541
Qy      554 LeuAlaPheGlyGlyTyLeuSerGluValValGluGluAsnThrProProLysMetGlu 573
Db      542 TTAGCATTTGGAGGCTACTTAAAGTGAAGTAGAAGAAATACTCCACCAAAATGGAA 601
Qy      574 LysGluGlyLeuGluIleMet 580
Db      602 AAAGAAGGCTTGGACATAATG 622

```

## RESULT 10

```

CN335014
LOCUS      17000531936046 GRN_EB Homo sapiens cDNA 5', mRNA sequence.
DEFINITION
ACCESSION  CN335014
VERSION     CN335014.1 GI:47334948
KEYWORDS    EST.
SOURCE      Homo sapiens
ORGANISM    Homo sapiens

```

## REFERENCE

```

AUTHORS    Brandenberger, R., Wei, H., Zhang, S., Lei, S., Murage, J., Fisk, G.J.,
            Li, Y., Xu, C., Fang, R., Guegler, K., Rao, M.S., Mandalam, R.,
            Lebkowski, J. and Stanton, L.W.

```

```

TITLE      Transcriptional characterization elucidates signaling networks that
            control human ES cell growth and differentiation
JOURNAL     Nat. Biotechnol. 22 (6), 707-716 (2004)
COMMENT     Contact: Brandenberger R
            Regenerative Medicine
            Geron Corporation
            230 Constitution Drive, Menlo Park, CA 94025, USA
            Tel: 650 473 8658
            Fax: 650 473 7760
            Email: rbrandenberger@geron.com
            Insert Length: 624 Std Error: 0.00.

```

```

FEATURES   Location/Qualifiers
            source          1..624
                        /organism="Homo sapiens"
                        /mol_type="mRNA"
                        /db_xref="taxon:9606"
                        /tissue_type="embryonic stem cells, embryoid bodies
                        derived from H1, H7 and H9 cells"
                        /clone_lib="GRN_EB"
                        /note="Oligo dr primed, full-length enriched cDNA library
                        from embryoid body outgrowths derived from hES cell lines
                        H1 (p32), H7 (p29), and H9 (p26) maintained in feeder-free
                        conditions."

```

```

ORIGIN
Alignment Scores:
Pred. No.:      4.11e-215      Length:      624
Score:          207.00      Matches:      207

```

```

This is the 5' sequence of the clone insert. Clone from S. Wiemann,
Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany

```



NCBI trace archive.  
 Plate: 32 row: E column: 07  
 Seq primer: SP6 primer  
 High quality sequence start: 61  
 High quality sequence stop: 283.  
 Location/Qualifiers  
 1..596  
 /organism="Sus scrofa"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9823"  
 /clone="PPSUBLIB 32E07"  
 /tissue\_type="Peyer's patch"  
 /clone\_lib="UNMMP3"

note="Organ: small intestine, jejunum; Vector: pCMVSPORT6; Site 1: NotI; Site 2: SalI; Jejunal Peyer's patches were isolated from healthy, 4-6 week old cross-bred pigs. RNA was extracted either immediately or after 3 hours stimulation in an Ussing chamber with one of the following treatments: 1) no treatment, 2) Salmonella choleraesuis vaccine strain SC-54, 3) phorbol myristate acetate, concanavalin A, and 8-bromo-cyclic AMP, and 4) lipopolysaccharide and cholera toxin. Each treatment was performed in the presence and absence of cycloheximide. Purified poly A + RNA from each of the treated tissues (2-4) was combined together, reverse transcribed, and cloned in to pCMVSPORT6 to make a library of approximately 530,000 recombinant clones with an average insert size of 1.0 kbp. Poly A+ RNA from freshly isolated and non-treated tissue from an Ussing chamber (treatment 1) was cloned in the same manner to produce an unstimulated cDNA library of approximately 900,000 clones with an average insert size of 1.5 kbp. Equal portions of the two libraries were pooled and then subtracted with porcine ST fibroblast RNA to create a subtracted porcine Peyer's patch library of approximately 6,000 clones with an average insert size of 1.0 kbp."

FEATURES  
 source  
 1..596  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="DMC02192"  
 /clone\_lib="Sugano cDNA library, dermoid cancer"  
 /note="dermoid cancer"

ORIGIN  
 Alignment Scores:  
 Pred. No.: 3,498-204 Length: 596  
 Score: 197.00 Matches: 197  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 20.76% Indels: 0  
 DB: 6 Gaps: 0

US-09-671-687A-3 (1-949) x CB286853 (1-596)

Qy 549 GluArgCysAsnSerLeuAlaPheGlyGlyTyrLeuSerGluValValGluGluAsnThr 568  
 Db 4 GAACGCTGTAACCTTTAGCACTTTGGAGGCTACTTAAAGTAGTAGTAGAGAAATACT 63  
 Qy 569 ProProLysMetGluLysGluGlyLeuGluLeuMetIleGlyLysLysGlyLeuGln 588  
 Db 64 CCACCGAAATGGAAGAGAGGTTTAGAGATAATGATTGGAAGAAAGAGATCCAG 123  
 Qy 589 GlyHisTyrAsnSerCysTyrLeuAspSerThrLeuPheCysLeuPheAlaPheSerSer 608  
 Db 124 GGTCAATTAACATCTGTACTTACCACTCAACCTTATCTGCTTATTCCTTTTATGTTCC 183  
 Qy 609 ValLeuAspThrValLeuLeuArgProLysGluLysAsnAspValGluTyrTyrSerGlu 628  
 Db 184 GTCTCGACACTGTCTTACTAGACCAAGAAAGAAAGATGATGATATATTATAGTGAA 243  
 Qy 629 ThrGlnGluLeuLeuArgThrGluLeuValAsnProLeuArgIleTyrGlyTyrValCys 648  
 Db 244 ACTCAAGAGCTACTGAGGACAGAAATGTTAATCTCTCCTGAGAATATACGATATGATCT 303  
 Qy 649 AlaThrLysIleMetLysLeuArgLysIleLeuGluLysValGluAlaAlaSerGlyPhe 668  
 Db 304 GCAACAAAGATTATGAACTGAGGAAATACTTGAAGAAAGTTGAGGCTGCATCAGGATTT 363  
 Qy 669 ThrSerGluGluLysAspProGluGluPheLeuAsnIleLeuPheHisIleLeuArg 688

Db 364 ACCCTCTGAAGAAAAAGATCCTCGAAGAAATTTTGAATATCCTGTTTCATCATATTTTAAGG 423  
 Qy 689 ValGluProLeuLeuLysIleArgSerAlaGlyGlnLysValGlnAspCysTyrPheTyr 708  
 Db 424 GTAGAACCATTTGTTAAATAAAGATCAGCAGGTCAAAAGATACAAGATTGTTACTTCTAT 483  
 Qy 709 GlnIlePheMetGluLysAsnGluLysValGlyValProThrIleGlnLeuLeuGlu 728  
 Db 484 CAAATTTTATGGAAGAAAAATGAGAAAGTTGAGTCTCTACAATTCAGCAGTTGTTAGAA 543  
 Qy 729 TrpSerPheIleAsnSerAsnLeuLysPheAlaGluAlaProSerCysLeu 745  
 Db 544 TGGTCTTTTATCAACAGTAACCTGAATTTTGGAGGACCACCATCATGCTG 594

RESULT 12  
 BP226285  
 LOCUS BP226285 Sugano cDNA library, dermoid cancer Homo sapiens cDNA  
 DEFINITION BP226285  
 ACCESSION BP226285  
 VERSION BP226285.1 GI:52099190  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 575)  
 AUTHORS Suzuki,Y., Yamashita,R., Shiota,M., Sakakibara,Y., Chiba,J.,  
 Mizushima,Sugano,J., Nakai,K. and Sugano,S.  
 TITLE Sequence comparison of human and mouse genes reveals a homologous  
 block structure in the promoter regions  
 JOURNAL Genome Res. 14 (9), 1711-1718 (2004)  
 COMMENT Contact: Yutaka Suzuki  
 Department of Virology  
 Institute of Medical Science, University of Tokyo  
 4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan  
 Email: ysuzuki@ims.u-tokyo.ac.jp.

FEATURES  
 source  
 1..575  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="DMC02192"  
 /clone\_lib="Sugano cDNA library, dermoid cancer"  
 /note="dermoid cancer"

ORIGIN  
 Alignment Scores:  
 Pred. No.: 1,948-195 Length: 575  
 Score: 189.00 Matches: 189  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 19.92% Indels: 0  
 DB: 5 Gaps: 0

US-09-671-687A-3 (1-949) x BP226285 (1-575)

Qy 351 LeuAsnGlySerSerValAspSerGlnProGlnSerLysSerLysAsnThrTrpTyrIle 370  
 Db 8 CTTAATGGTCTCTCTGTTGACTCAACACCACCAATCCAAATCAAAATAATCATGGTACATT 67  
 Qy 371 AspGluValAlaGluAspProAlaLysSerLeuThrGluIleSerThrAspPheAspArg 390  
 Db 68 GATGAAGTTGCAAGAGACCCCTGCAAAATCTCTTACAGAGATATCTACAGACTTTGACCGT 127  
 Qy 391 SerSerProProLeuGlnProProValAsnSerLeuThrGluAsnArgPheHis 410  
 Db 128 TCTTCACCACTCACTCAGCTCTCTCTGTGAATCTCTTACAGAGATATCTACAGACTTTGACCGT 187  
 Qy 411 SerLeuProPheSerLeuThrLysMetProAsnThrAsnGlySerIleGlyHisSerPro 430  
 Db 188 TCTTTTACCATTTCAGTCTCACCAGATGCCAATACCAATGGAAGTATTGCGCCACAGTCCA 247

QY 431 LeuSerLeuSerAlaGlnSerValMetGluGluLeuAasnThrAlaProValGlnGluSer 450  
Db 248 CTTTCTCTGTAGCCAGCTGTATATGAAGAGCTAAACACATGCACCGCTCCAGAGAGT 307  
QY 451 ProProLeuAlaMetProProGlyAasnSerHisGlyLeuGluValGlySerLeuAlaGlu 470  
Db 308 CCACCCCTGGCCATGCTCTCTGGGAACCTACATGGTCTAGAAGTGGGCTCATTTGGCTGA 367  
QY 471 ValLysGluAasnProProPheTyrGlyValIleArgTyrIleGlyGlnProProGlyLeu 490  
Db 368 GTTAGAGGAACCCCTCTTCATGGGGTAATCCGTGGATCGGTGAGCCACGAGACTG 427  
QY 491 AasnGluValLeuAlaGlyLeuGluLeuGluAaspGlyCysAlaGlyCysThrAaspGlyThr 510  
Db 428 AATGAAGTCTGCTGGACTGGACTGGACTGGAGATGAGTGTGAGGCTGTACGGATGGACC 487  
QY 511 PheArgGlyThrArgTyrPheThrCysAlaLeuLysGlyAlaLeuPheValLysLeuLys 530  
Db 488 TTACAGAGGCACTCGGTATTTACCTGTGCCCTGAAGAAGCGCGTGTGTGAAACTGAAG 547  
QY 531 SerCysArgProAaspSerArgPheAla 539  
Db 548 AGCTGAGGCCCTGACTCTAGGTTTGCA 574

## RESULT 13

CF528626/c  
LOCUS  
DEFINITION  
UI-1-BC1-ajd-e-08-0-UI.81 NCI CGAP\_P12 Homo sapiens cDNA clone  
UI-1-BC1-ajd-e-08-0-UI.3', mRNA sequence.

## ACCESSION

VERSION

## KEYWORDS

SOURCE

## ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 581)

NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

Unpublished (1997)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Dr. Steven Brown

cDNA Library preparation: Dr. M. Bento Soares, University of Iowa

cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: Clone distribution information can be obtained

from Dr. M. Bento Soares, bento-soares@uiowa.edu

Seq primer: M13 FORWARD

POLYA=Yes.

## FEATURES

source

Location/Qualifiers  
1..581  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="UI-1-BC1-ajd-e-08-0-UI"  
/tissue\_type="Placenta"  
/dev\_stage="8-9 weeks"  
/lab\_host="DH10B (Life Technologies)"  
/clone\_lib="NCI CGAP\_P12"  
/note="Organ: Placenta; Vector: pTT73-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; NCI CGAP\_P12 is a subcloned cDNA library constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pTT73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The

sequence tags for this library are GA, AGGA. For additional information, contact: Bento Soares, bento-soares@uiowa.edu  
TAG TISSUE=placenta human 8 week  
TAG\_LIB=UI-1-BC1  
TAG\_SEQ=GA"

## ORIGIN

Alignment Scores:  
Pred. No.: 1,96e-195 Length: 581  
Score: 189.00 Matches: 189  
Percent Similarity: 100.00% Conservations: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 19.92% Indels: 0  
DB: 7 Gaps: 0  
US-09-671-687A-3 (1-949) x CF528626 (1-581)  
QY 526 PheValLysLeuLysSerCysArgProAaspSerArgPheAlaSerLeuGlnProValSer 545  
Db 579 TTTGTGAACCTGAAGAGCTGCAGCGCTGACTCTAGGTTTCATCATTTGCAGCGGTTTCC 520  
QY 546 AasnGlnIleGluArgCysAasnSerLeuAlaPheGlyGlyTyrLeuSerGluValGlu 565  
Db 519 AATCAGATTGAGCGCTGTAACTCTTTAGCATTTTGGAGGCTACTTAAGTAGTAGTAA 460  
QY 566 GluAasnThrProProLysMetGluLysGluGluLysLeuMetIleGlyLysLysLys 585  
Db 459 GAAATATCTCCCAAAATGGAAGAAAGAGCTTGGAGATATATGATTGGAGAGAGAA 400  
QY 586 GlyIleGlnGlyHisTyrAasnSerCysTyrLeuAaspSerThrLeuPheCysLeuPheAla 605  
Db 399 GGCATCCAGGCTCATTAATCTTGTACTTAGACTCAACCTATTCTGCTATTGCT 340  
QY 606 PheSerSerValLeuAaspThrValLeuLeuAurProLysGluLysAasnAaspValGluTyr 625  
Db 339 TTTAGTTCTGTTCTGGACACTGTGTACTTAGACCCCAAGAAAGAACGATGAGATAT 280  
QY 626 TyrSerGluThrGlnGluLeuLeuArgThrGluIleValAasnProLeuArgIleTyrGly 645  
Db 279 TATAGTGAACCCAGAGCTACTGAGGACAGAAATTTGTTAATCTCTCGAATATATGGA 220  
QY 646 TyrValCysAlaThrLysIleMetLysLeuArgLysIleLeuGluLysValGluAlaLa 665  
Db 219 TATGTGTGTGCCACAAAATTTATGAACCTGAGAAATATCTTGAAGAGTGGAGCTGCA 160  
QY 666 SerGlyPheThrSerGluLysAaspProGluGluPheLeuAasnIleLeuPheHis 685  
Db 159 TCAGGATTTACCTCTCAAGAAAAGATCCTGAGGAATTTCTTGAATATTCTGTTTCATCAT 100  
QY 686 IleLeuArgValGluProLeuLeuLysIleArgSerAlaGlyGlnLysValGlnAaspCys 705  
Db 99 ATTTTAAGGGTAGAACCTTTGCTAAAAATAAGATCAGCAGTCAAAAGGTACAAGATTGT 40  
QY 706 TyrPheTyrGlnIlePheMetGluLys 714  
Db 39 TACTTCTATCAAAATTTTATGGAAGAAA 13

## RESULT 14

LOCUS

CD641429

DEFINITION

AGENCOURT\_14532010 NIH\_MGC\_191 Homo sapiens cDNA clone

IMAGE:30418040 5', mRNA sequence.

ACCESSION

CD641429

VERSION

CD641429.1 GI:31810295

KEYWORDS

EST.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 736)

NIH-MGC http://mhc.nci.nih.gov/.

AUTHORS

National Institutes of Health, Mammalian Gene Collection (MGC)

CD641429 736 bp mRNA linear EST 17-JUN-2003  
AGENCOURT\_14532010 NIH\_MGC\_191 Homo sapiens cDNA clone  
IMAGE:30418040 5', mRNA sequence.

ACCESSION  
CD641429  
VERSION  
CD641429.1 GI:31810295  
KEYWORDS  
EST.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 736)

NIH-MGC http://mhc.nci.nih.gov/.

AUTHORS

National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL  
COMMENT

Unpublished (1999)  
Contact: Daniela S. Gerhard, Ph.D.  
Office of Cancer Genomics  
National Cancer Institute / NIH  
Bldg. 31 Rm10A07 Bethesda, MD 20892  
Email: c9apbs-remail.nih.gov  
Tissue Procurement: Narayan Bhat  
cDNA Library Preparation: CLONTECH Laboratories, Inc.  
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: NDCM220 row: 9 column: 09  
High quality sequence stop: 594.  
Location/Qualifiers

FEATURES  
source

1..736  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:30418040"  
/tissue\_type="Pooled"  
/lab\_host="DH10B (T1 phage-resistant)"  
/clone\_lib="NIH\_MGC\_191"  
/note="Vector: pDNR-LIB; Site 1: SfiI (ggccattatggcc);  
Site 2: SfiI (ggccgctctggcc); Library is oligo-dT primed  
and directionally cloned. BMC - peripheral Blood  
Mononuclear Cells. RNA was pooled from 3/6hour stimulation  
with PMA and ionomycin. 5' and 3' adaptors were used in  
cloning as follows: 5' adaptor sequence:  
5'-CACGCCATTATGGC-3' and 3' adaptor sequence:  
5'-ATTCAGAGCGCGCGCCGACATG-dt(30)BN-3' (where B = A,  
C, or G and N = A, C, G, or T). Average insert size 1.69  
kb (range 0.70-5.0 kb). 15/15 colonies contained inserts  
by PCR. This library was enriched for full-length clones  
and was constructed by Clontech Laboratories (Palo Alto,  
CA). Note: this is a NIH\_MGC Library."

## ORIGIN

Alignment Scores:  
Pred. No.: 5.98e-191 Length: 736  
Score: 185.00 Matches: 185  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 19.49% Indels: 0  
DB: 6 Gaps: 0

US-09-671-687A-3 (1-949) x CD641429 (1-736)

Qy 458 GlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLysGluAsnProPhe 477  
Db 1 CGGAACTCACATGGTCTAGAGTGGGCTCATGGCTGAAGTAAAGAGAACCTCTTTC 60  
Qy 478 TyrGlyValIleArgTrpIleGlyGlnProGlyLeuAsnGluValLeuAlaGlyLeu 497  
Db 61 TATGGGTAACTCGTTGGATCGGTAGCCAGCCAGGACTGAATGAAGTCTCGCTGACTG 120  
Qy 498 GluLeuGluAspGluCysAlaGlyCysThrAspGlyThrPheArgGlyThrArgTyrPhe 517  
Db 121 GAATCGAAGATGAGTGTGACGGCTGTACGGATGAACCTTCAGAGGCACCTCGGTATTTC 180  
Qy 518 ThrCysAlaLeuLysLeuAlaLeuPheValLysLeuLysSerCysArgProAspSerArg 537  
Db 181 ACCTGTGCGCCCTGAAGAAGCGCTGTGTGTGAACCTGAAGAGCTGAGGCGCTGACTTAGG 240  
Qy 538 PheAlaSerLeuGlnProValSerAsnGlnIleGluArgCysAsnSerLeuAlaPheGly 557  
Db 241 TTTCATCATTTGCACCGCTTTCATCATTTGCATTTGCACTTTTGGCATTTGA 300  
Qy 558 GlyTyrLeuSerGluValValGluAlaGluAsnThrProProLysMetGluLysGluGlyLeu 577  
Db 301 GGCTACTTAAGTGAAGTAGTAGAAGAAATACTCCACCAAAATGGAAAAAGAGCGTTG 360

Qy 578 GluLeuMetIleGlyLysGlyLysGlyLeuGlnGlyHisTyrAsnSerCysTyrLeuAsp 597  
Db 361 GAGATAATGATTGGGAAGAGAAAGCATCCAGGGTCATTACAAATCTCTTTACTTAGAC 420  
Qy 598 SerThrLeuPheCysLeuPheAlaPheSerSerValLeuAspThrValLeuLeuArgPro 617  
Db 421 TCACCTTATTCTGCTTATTCTTTAGTTCTGTCTGGACACTGTGTACTTAGACCC 480  
Qy 618 LysGluLysAsnAspValGluTyrTyrSerGluThrGlnGluLeuLeuArgThrGluLeu 637  
Db 491 AAAGAAAAGACGATGATGATATTATATGAAACCCAGAGCTACTGAGGACAGAAATT 540  
Qy 638 ValAsnProLeuArg 642  
Db 541 GTTAATCCTCTGAGA 555

## RESULT 15

BG180100

LOCUS

DEFINITION

602329675F1 NIH\_MGC\_91 Homo sapiens cDNA clone IMAGE:4430971 5',

mRNA sequence.

ACCESSION

BG180100

VERSION

BG180100.1 GI:12686803

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

1 (bases 1 to 919)

AUTHORS

NIH-MGC http://mgc.nci.nih.gov/.

JOURNAL

National Institutes of Health, Mammalian Gene Collection (MGC)

COMMENT

Contact: Robert Strausberg, Ph.D.

Email: c9apbs-remail.nih.gov

Tissue Procurement: DCTD/DTF

cDNA Library Preparation: Life Technologies, Inc.

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLAM10185 row: e column: 20

High quality sequence stop: 731.

Location/Qualifiers

1..919

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:4430971"

/tissue\_type="adenocarcinoma, cell line"

/lab\_host="DH10B (phage-resistant)"

/clone\_lib="NIH\_MGC\_91"

/note="Organ: prostate; Vector: pCMV-SPORT6; Site 1: NotI;

Site 2: SalI; Cloned unidirectionally; oligo-dT primed.

Average insert size 1.4 kb. Library enriched for

full-length clones and constructed by Life Technologies.

Note: this is a NIH\_MGC Library."

## ORIGIN

Alignment Scores:

Pred. No.: 2.79e-184 Length: 919

Score: 179.00 Matches: 179

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 18.86% Indels: 0

DB: 4 Gaps: 0

US-09-671-687A-3 (1-949) x BG180100 (1-919)

Qy 771 ThrAspLeuLeuGluAspThrProArgGlnCysArgIleCysGlyGlyLeuAlaMetTyr 790

Db 2 ACAGATTACTTGAAGACACTCCCAAGACAGTGGCCGGATATGTGGAGGCGCTTGCATGTAT 61



```
QY 791 GluCysArgGluCysTyrAspAspProAspIleSerAlaGlyLysIleLysGlnPheCys 810
Db |||||
62 GAGTGTAGAGAAATGCTACGACGATCCGGACATCTCAGCTGGAAAAATCAAGCAGTTTGT 121
QY 811 LysThrCysAsnThrGlnValHisLeuHisProLysArgLeuAsnHisLysTyrAsnPro 830
Db |||||
122 AAAACCTGCAACACTCAAGTCCACCTTCATCCGAGAGGCTGAATCAATAATAACCCA 181
QY 831 ValSerLeuProLysAspLeuProAspTyrAspTyrArgHisGlyCysIleProCysGln 850
Db |||||
182 GTGTCACTTCCCAGAGACTTACCCGACTGGGACTGGAGACACGGCTGCATCCCTTGCCAG 241
QY 851 AsnMetGluLeuPheAlaValLeuCysIleGluThrSerHisTyrValAlaPheValLys 870
Db |||||
242 AATATGGAGTTATTGCTGCTTCTGTCATAGAAACAAGCCACTATGTTGCTTTTGTGAAG 301
QY 871 TyrGlyLysAspAspSerAlaTyrLeuPhePheAspSerMetAlaAspArgAspGlyGly 890
Db |||||
302 TATGGGAAGGACGATTTCTGCTGGCTCTTCTTTGACAGCATGGCCGATCGGATGGTGGT 361
QY 891 GlnAsnGlyPheAsnIleProGlnValThrProCysProGluValGlyGluTyrLeuLys 910
Db |||||
362 CAGAAATGGCTTCAACATTCCTCAAGTCACCCCATGCCAGAGTAGGAGAGTACTTGAAG 421
QY 911 MetSerLeuGluAspLeuHisSerLeuAspSerArgArgIleGlnGlyCysAlaArgArg 930
Db |||||
422 ATGTCTCTGGAGACCTGCATTCTCTGGACTCCAGGAGATCCAGGCTGTGCACCAAGA 481
QY 931 LeuLeuCysAspAlaTyrMetCysMetTyrGlnSerProThrMetSerLeuTyrLys 949
Db |||||
482 CTGCTTTGTGATGATATATGTGATGTACAGAGTCCAAACAAATGAGTTTGTACAAA 538
```

Search completed: April 17, 2005, 11:37:04  
Job time : 6214 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 18, 2005, 09:17:58 ; Search time 22 Seconds  
(without alignments)  
4150.441 Million cell updates/sec

Title: US-09-671-687A-3  
Perfect score: 949  
Sequence: 1 MSSGLWSQEKVTSPYWEERI.....RLLCDAYMCVMQSPMTSLYK 949

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 283416 seqs, 96216763 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : PIR\_79:.\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	12	1.3	727	2 S42834	F40F12.5 protein -
2	12	1.3	1021	2 F88568	protein F40F12.5 [
3	8	0.8	150	2 S65996	transcription regu
4	8	0.8	215	2 C83149	probable transcrip
5	8	0.8	448	2 T15188	hypothetical prote
6	8	0.8	450	2 AG1370	glucose-6-phosphat
7	8	0.8	450	2 AE1740	glucose-6-phosphat
8	8	0.8	458	2 D88950	protein R09B5.1 [i
9	8	0.8	486	2 C64765	yait protein precu
10	8	0.8	524	2 JS0746	cytochrome P450 1A
11	8	0.8	551	2 S63361	probable membrane
12	8	0.8	651	2 G83395	probable AMP-bind
13	8	0.8	740	2 E69420	hydrogenase expres
14	8	0.8	760	2 F75530	ribonuclease - Del
15	8	0.8	1086	2 T33893	hypothetical prote
16	7	0.7	1056	2 A71570	hypothetical prote
17	7	0.7	72	2 G97134	hypothetical prote
18	7	0.7	81	2 C97140	hypothetical prote
19	7	0.7	104	2 T13628	hypothetical prote
20	7	0.7	116	2 I65342	tachykinin B precu
21	7	0.7	116	2 A43779	neurokinin B precu
22	7	0.7	123	2 T09268	probable tail comp
23	7	0.7	126	2 C82665	hypothetical prote
24	7	0.7	143	2 D71911	ribonuclease hi -
25	7	0.7	152	2 A10448	probable ribonucle
26	7	0.7	152	2 A84303	hypothetical prote
27	7	0.7	157	2 S68012	probable olfactory
28	7	0.7	159	2 AD1590	hypothetical prote
29	7	0.7	163	2 JC5045	epithelial membran

30	7	0.7	169	2 C91130	hypothetical prote
31	7	0.7	169	2 C85975	hypothetical prote
32	7	0.7	169	2 A65103	pts system, n-acet
33	7	0.7	171	2 E95002	conserved hypothet
34	7	0.7	174	2 A45356	neuromedin U precu
35	7	0.7	177	2 F96739	hypothetical prote
36	7	0.7	178	2 A75578	transcription regu
37	7	0.7	179	2 AF1362	hypothetical prote
38	7	0.7	180	2 A69387	fumarase (fum-1) h
39	7	0.7	189	2 F90780	probable antirepre
40	7	0.7	189	2 E90820	antirepressor prot
41	7	0.7	189	2 A90852	antirepressor prot
42	7	0.7	189	2 E90874	probable antirepre
43	7	0.7	189	2 A90902	probable antirepre
44	7	0.7	189	2 B90911	probable antirepre
45	7	0.7	189	2 G90999	probable antirepre

ALIGNMENTS

RESULT 1

S42834  
F40F12.5 protein - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 24-Nov-1999  
C:Accession: S42834  
R:Kershaw, J.  
submitted to the EMBL Data Library, February 1994  
A:Reference number: S42830  
A:Accession: S42834  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-727 <KER>  
A:Cross-references: EMBL:X30215  
C:Genetics:  
A:Introns: 19/1; 53/2; 106/3; 156/2; 198/1; 358/2; 441/2; 678/3  
C:Superfamily: Caenorhabditis elegans F40F12.5 protein

Query Match 1.3%; Score 12; DB 2; Length 727;  
Best Local Similarity 100.0%; Pred. No. 0.0017;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 856 AVLCTSTSHVA 867  
Db 592 AVLCTSTSHVA 603

RESULT 2

F88568  
protein F40F12.5 [imported] - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 10-May-2001  
C:Accession: F88568  
R:Anonymous, The C. elegans Sequencing Consortium.  
Science 282, 2012-2018, 1998  
A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biolog  
A:Reference number: A75000; MUID:99069613; PMID:9851916  
A>Note: see websites genome.wustl.edu/gsc/C\_elegans/ and www.sanger.ac.uk/projects/C\_ele  
A>Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and  
A:Accession: F88568  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-1021 <STO>  
A:Cross-references: GB:chr\_III; PIDN:CAA82939.1; PID:g3877002; GSPDB:GN00021; CESP:F40F1

Query Match 1.3%; Score 12; DB 2; Length 1021;  
Best Local Similarity 100.0%; Pred. No. 0.0023;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 856 AVLCTSHYVA 867  
 |||||  
 Db 886 AVLCTSHYVA 897

## RESULT 3

S65996

transcription regulator homolog yybA - Bacillus subtilis

C/Species: Bacillus subtilis

C/Date: 28-Oct-1996 #sequence\_revision 13-Mar-1997 #text\_change 09-Jul-2004

C/Accession: S65996; D70086

R/Ogasawara, N.; Nakai, S.; Yoshikawa, H.

DNA Res. 1, 1-14, 1994

A/Title: Systematic sequencing of the 180 kilobase region of the Bacillus subtilis chrom

A/Reference number: S65967; MUID:96051385; PMID:7584024

A/Accession: S65996

A/Status: preliminary; nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Residues: 1-150 &lt;OGA&gt;

A/Cross-references: UNIPROT:P37503; EMBL:D26185; NID:g467326; PIDN:BAA05202.1; PID:dl005

A/Note: the nucleotide sequence was submitted to the EMBL Data Library, December 1993

R/Kunst, P.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte

C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch

A.; Ehrlich, S.D.; Emmeron, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.

Nature 390, 249-256, 1997

A/Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gall

iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.

Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,

A/Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel

Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle

Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanlon,

A/Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seron

akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpetra, P.; Tognoni, A.; Tosato, V.; Uchiyama

T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yanane, K.; Yasumoto, K.; Yata, K.; Yoshida, Y

A/Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.

A/Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.

A/Reference number: A69580; MUID:98044033; PMID:9384377

A/Accession: D70086

A/Status: preliminary; nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Residues: 1-150 &lt;KUN&gt;

A/Cross-references: GB:Z99124; GB:AL009126; NID:g2636442; PIDN:CAB16108.1; PID:el184797;

A/Experimental source: strain 168

C/Genetics:

A/Gene: yybA

Query Match

Best Local Similarity 0.8%; Score 8; DB 2; Length 150;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 666 SGPTSEK 673

|||||

Db 114 SGPTSEK 121

## RESULT 4

C83149

probable transcription regulator PA3973 [imported] - Pseudomonas aeruginosa (strain PA01

C/Species: Pseudomonas aeruginosa

C/Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 09-Jul-2004

C/Accession: C83149

R/Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Br

adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lardig, K.; Lim,

; Lory, S.; Olson, M.V.

Nature 406, 959-964, 2000

A/Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho

A/Reference number: A82950; MUID:20437337; PMID:10984043

A/Accession: C83149

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-215 &lt;STO&gt;

A/Cross-references: UNIPROT:Q9HX43; GB:AE004815; GB:AE004091; NID:g9950159; PIDN:AAG0736

A/Experimental source: strain PA01

C/Genetics:

A/Gene: PA3973

Query Match

Best Local Similarity 0.8%; Score 8; DB 2; Length 215;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 164 VELLEGR 171

|||||

Db 138 VELLEGR 145

## RESULT 5

T15188

hypothetical protein F55A12.6 - Caenorhabditis elegans

C/Species: Caenorhabditis elegans

C/Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 09-Jul-2004

C/Accession: T15188

R/Pauley, A.; Gattung, S.

submitted to the EMBL Data Library, May 1997

A/Description: The sequence of C. elegans cosmid F55A12.

A/Reference number: Z18305

A/Accession: T15188

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-448 &lt;PAU&gt;

A/Cross-references: UNIPROT:O01754; EMBL:AF003130; NID:g2088663; PID:g2088666; PIDN:AAR5

A/Experimental source: strain Bristol N2; clone F55A12

C/Genetics:

A/Gene: CESP:F55A12.6

A/Map position: 1

A/Introns: 40/1, 98/2, 198/1, 232/1, 251/2, 370/2, 409/1

Query Match

Best Local Similarity 0.8%; Score 8; DB 2; Length 448;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 540 SLQPVSNQ 547

|||||

Db 85 SLQPVSNQ 92

## RESULT 6

AG1370

Glucose-6-phosphate isomerase [imported] - Listeria monocytogenes (strain EGD-e)

C/Species: Listeria monocytogenes

C/Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 09-Jul-2004

C/Accession: AG1370

R/Glaeser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker

; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.

D.; Jones, L.M.; Karst, U.

Science 294, 849-852, 2001

A/Authors: Krefte, J.; Kuhn, M.; Kunat, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma

ok, C.; Schluteter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,

A/Title: Comparative genomics of Listeria species

A/Reference number: AB1077; MUID:21537279; PMID:11679669

A/Accession: AG1370

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-450 &lt;GLA&gt;

A/Cross-references: UNIPROT:O8Y4R7; GB:NC\_003210; PIDN:CAD00445.1; PID:g16411855; GSPDB:

A/Experimental source: strain EGD-e

C/Genetics:

A/Gene: pgi

C/Superfamily: glucose-6-phosphate isomerase

Query Match

Best Local Similarity 0.8%; Score 8; DB 2; Length 450;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 45 SIGQYIQD 52

|||||

Db 312 SIGQYIQD 319

```

RESULT 7
AE1740
glucose-6-phosphate isomerase [imported] - Listeria innocua (strain Clip11262)
C;Species: Listeria innocua
C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C;Accession: AE1740
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker,
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A;Authors: Krett, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma
ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
A;Title: Comparative genomics of Listeria species.
A;Reference number: AE1077; MUID:21537279; PMID:11679669
A;Accession: AE1740
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-450 <GLA>
A;Cross-references: UNIPROT:Q92886; GB:AL592022; PIDN:CAC97693.1; PID:g16414988; GSPDB:C
A;Experimental source: strain Clip11262
C;Genetics:
A;Gene: pgi
C;Superfamily: glucose-6-phosphate isomerase

Query Match 0.8%; Score 8; DB 2; Length 450;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 45 SIGQYIQD 52
|||||||
DB 312 SIGQYIQD 319

RESULT 8
D88950
protein R09B5.1 [imported] - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C;Accession: D88950
R;anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A;Title: Genome sequence of the nematode C. elegans: a platform for investigating biolog
A;Reference number: A75000; MUID:99069613; PMID:9851916
A;Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/Projects/C_ele
A;Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
A;Accession: D88950
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-458 <STO>
A;Cross-references: UNIPROT:O44617; GB:chr_V; PIDN:AAB94210.1; PID:g2736431; GSPDB:GN000
C;Genetics:
A;Gene: R09B5.1
A;Map position: 5

Query Match 0.8%; Score 8; DB 2; Length 458;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 759 FKXIFPSL 766
|||||||
DB 90 FKXIFPSL 97

RESULT 9
C64765
yair protein precursor - Escherichia coli (strain K-12)
C;Species: Escherichia coli
C;Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 09-Jul-2004
C;Accession: C64765
R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
A.; Rose, D.J.; Mau, B.; Shao, Y.

```

```

Science 277, 1453-1462, 1997
A;Title: The complete genome sequence of Escherichia coli K-12.
A;Reference number: A64720; MUID:97426617; PMID:9278503
A;Accession: C64765
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-486 <BLAT>
A;Cross-references: UNIPROT:P77199; GB:AE000144; GB:U00096; NID:g1786568; PIDN:AAC73474
A;Experimental source: strain K-12, substrain MG1655
C;Genetics:
A;Gene: yair
F;29-486/Product: yair protein #status predicted <SIG>
F;29-486/Product: yair protein #status predicted <MAT>

Query Match 0.8%; Score 8; DB 2; Length 486;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 212 SDYAGFGD 219
|||||||
DB 273 SDYAGFGD 280

RESULT 10
JS0746
cytochrome P450 1A1 - Chinese hamster
N;Contains: oxidoreductase (EC 1.-.-.-)
C;Species: Cricetulus griseus (Chinese hamster)
C;Date: 09-Oct-1992 #sequence_revision 09-Oct-1992 #text_change 20-Jun-2000
C;Accession: JS0746
R;Ohgiya, S.; Goda, T.; Ishizaki, K.; Morimoto, M.; Sakamoto, T.; Kamataki, T.; Shinriki
submitted to JIPID, September 1992
A;Reference number: JS0746
A;Accession: JS0746
A;Molecule type: mRNA
A;Residues: 1-524 <OHG>
A;Cross-references: DDBJ:D12977; NID:g220307; PIDN:BAA02354.1; PID:g220308
A;Note: this is a revision to the sequence from reference JX0189
A;Note: the source is designated as golden hamster
C;Superfamily: human cytochrome P450 CYP2D6; cytochrome P450 homology
C;Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; monooxygenase
F;318-483/Domain: cytochrome P450 homology <P45>
F;461/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 0.8%; Score 8; DB 2; Length 524;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 96 TELLALIT 103
|||||||
DB 14 TELLALIT 21

RESULT 11
S63361
probable membrane protein YNR030w - yeast (Saccharomyces cerevisiae)
N;Alternate names: hypothetical protein N3265
C;Species: Saccharomyces cerevisiae
C;Date: 27-Apr-1996 #sequence_revision 03-May-1996 #text_change 09-Jul-2004
C;Accession: S63361
R;Pohl, T.M.
submitted to the Protein Sequence Database, April 1996
A;Reference number: S63346
A;Accession: S63361
A;Molecule type: DNA
A;Residues: 1-551 <POH>
A;Cross-references: UNIPROT:P53730; EMBL:Z71645; NID:g1302524; PID:g1302525; MIPS:YNR030
C;Genetics:
A;Gene: SGD:ECM39
A;Cross-references: SGD:S0005313; MIPS:YNR030w
A;Map position: 14R
C;Keywords: transmembrane protein

```

F;3-19/Domain: transmembrane #status predicted <TM1>  
 F;56-72/Domain: transmembrane #status predicted <TM2>  
 F;141-157/Domain: transmembrane #status predicted <TM3>  
 F;166-182/Domain: transmembrane #status predicted <TM4>  
 F;187-203/Domain: transmembrane #status predicted <TM5>  
 F;210-226/Domain: transmembrane #status predicted <TM6>  
 F;308-324/Domain: transmembrane #status predicted <TM7>  
 F;332-348/Domain: transmembrane #status predicted <TM8>  
 F;363-379/Domain: transmembrane #status predicted <TM9>

Query Match 0.8%; Score 8; DB 2; Length 551;  
 Best Local Similarity 100.0%; Pred. No. 21;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 608 SVLDTVLL 615  
 |||||  
 Db 4 SVLDTVLL 11

## RESULT 12

G83395  
 C;Species: Pseudomonas aeruginosa (strain PAO1)  
 C;Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 09-Jul-2004  
 C;Accession: G83395  
 R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Boman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Latbig, K.; Lim, J.; Lory, S.; Olson, M.V.  
 Nature 406, 959-964, 2000  
 A;Title: Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic pathogen  
 A;Reference number: A82950; MUID:20437337; PMID:10984043  
 A;Accession: G83395  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-651 <STO>  
 A;Cross-references: UNIPROT:Q912B2; GB:AE004626; GB:AE004091; NID:g9947995; PIDN:AAG0538  
 A;Experimental source: strain PAO1  
 C;Genetics:  
 A;Gene: PA1997  
 F;135-619/Domain: acetate-CoA ligase homology <ACL>

Query Match 0.8%; Score 8; DB 2; Length 651;  
 Best Local Similarity 100.0%; Pred. No. 24;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 490 LNEVLGL 497  
 |||||  
 Db 209 LNEVLGL 216

## RESULT 13

E69420  
 C;Species: Archaeoglobus fulgidus  
 C;Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 14-Apr-2003  
 C;Accession: E69420  
 R;Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson, J.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.; Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.  
 Nature 390, 364-370, 1997  
 A;Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Attiach, P.; Kaine, B.P.; Sykes, S.; Smith, H.O.; Woese, C.R.; Venter, J.C.  
 A;Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeon  
 A;Reference number: A69250; MUID:98049343; PMID:9389475  
 A;Accession: E69420  
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A;Molecule type: DNA  
 A;Residues: 1-740 <KLE>  
 A;Cross-references: GB:AE001009; GB:AE000782; NID:g2689332; PIDN:AAB99876.1; PID:g264920  
 C;Superfamily: carbamoyl phosphate-converting enzyme ([NifH]-hydrogenase maturation factor)  
 Query Match 0.8%; Score 8; DB 2; Length 740;  
 Best Local Similarity 100.0%; Pred. No. 27;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 675 PEEFLNII 682  
 |||||  
 Db 48 PEEFLNII 55

## RESULT 14

F75530  
 C;Species: Deinococcus radiodurans (strain R1)  
 C;Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jul-2004  
 C;Accession: F75530  
 R;White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.; Ma, M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zaleski, C.; Smith, H.O.; Venter, J.C.; Fraser, C.M.  
 Science 286, 1571-1577, 1999  
 A;Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.  
 A;Reference number: A75250; MUID:20036996; PMID:10567266  
 A;Accession: F75530  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-760 <WHI>  
 A;Cross-references: UNIPROT:Q9RXG0; GB:AE001895; GB:AE000513; NID:g6458024; PIDN:AAF0993  
 A;Experimental source: strain R1  
 C;Genetics:  
 A;Gene: DR0353  
 A;Map position: 1  
 C;Superfamily: virulence-associated protein vacB homolog

Query Match 0.8%; Score 8; DB 2; Length 760;  
 Best Local Similarity 100.0%; Pred. No. 28;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 769 NITDLLED 776  
 |||||  
 Db 23 NITDLLED 30

## RESULT 15

T33893  
 C;Species: Caenorhabditis elegans  
 C;Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 09-Jul-2004  
 C;Accession: T33893  
 R;Maggi, L.  
 submitted to the EMBL Data Library, February 1999  
 A;Description: The sequence of C. elegans cosmid Y37E11B.  
 A;Reference number: Z21433  
 A;Accession: T33893  
 A;Status: preliminary; translated from GB/EMBL/DBJ  
 A;Molecule type: DNA  
 A;Residues: 1-1086 <MAG>  
 A;Cross-references: UNIPROT:Q9TYN3; EMBL:AF125451; PIDN:AAD12824.1; GSPDB:GNO00022; CESP:  
 A;Experimental source: strain Bristol N2; clone Y37E11B  
 C;Genetics:  
 A;Gene: CESP:Y37E11B.4  
 A;Map position: 4  
 A;Introns: 77/3; 305/3; 540/2; 729/2; 761/2; 981/2

Query Match 0.8%; Score 8; DB 2; Length 1086;  
 Best Local Similarity 100.0%; Pred. No. 38;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 393 PPLQPPPV 400  
 |||||  
 Db 46 PPLQPPPV 53

Search completed: April 18, 2005, 09:41:05  
 Job time : 25 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: April 17, 2005, 09:55:18 ; Search time 1145 Seconds  
(without alignments)  
5028.217 Million cell updates/sec

Title: US-09-671-687A-3  
Perfect score: 949  
Sequence: 1 MSSGLWSQKVTSPWEERI.....RLLCDAYMCWQPTMSLYK 949

Scoring table: OLIGO  
Xgapop 60.0 , Xgapext 60.0  
Ygapop 60.0 , Ygapext 60.0  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 5622541 seqs, 303355566 residues

Word size: 1  
Total number of hits satisfying chosen parameters: 11234615

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:  
-MODEL=frame+p2n.model -DEV=xlp  
-Q=/cgn2\_1/USPTO.spool.p/US09671687/runat\_15042005\_163040\_12059/app\_query.fasta\_1.1095  
-DB=Published Applications NA -QFMT=fastap -SUFFIX=Oligo.rnpb -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -SPART=1 -END=1 -MATRIX=oligo  
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=quality -THR\_MIN=1  
-ALIGN=15 -MODE=LOCAL -OUTFMT=ptco -NORM=ext -HEAFSIZE=500 -MINLEN=0  
-MAXLEN=2000000000 -USER=US09671687 @CGN 1.1 723 @runat\_15042005\_163040\_12059  
-NCPU=6 -ICPU=3 -NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100  
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : Published Applications NA:  
1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:  
2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:  
3: /cgn2\_6/ptodata/2/pubpna/US05\_NEW\_PUB.seq:  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:  
6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq:  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:  
9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq:  
10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq:  
11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq:  
12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:  
13: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq:  
14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq:  
15: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq:  
16: /cgn2\_6/ptodata/2/pubpna/US10D\_PUBCOMB.seq:  
17: /cgn2\_6/ptodata/2/pubpna/US10E\_PUBCOMB.seq:  
18: /cgn2\_6/ptodata/2/pubpna/US10F\_PUBCOMB.seq:  
19: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:  
20: /cgn2\_6/ptodata/2/pubpna/US11\_NEW\_PUB.seq:  
21: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:  
22: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	711	74.9	4527	18	US-10-921-707-25
2	685	72.2	5371	10	US-09-851-673-3
3	685	72.2	5371	18	US-10-755-889-489
4	641	67.5	2523	15	US-10-037-270-290
5	641	67.5	2523	17	US-10-117-722-290
6	279	29.4	1151	17	US-10-264-237-204
C 7	171	18.0	617	11	US-09-969-034-4253
C 8	165	17.4	2116	18	US-10-761-370-3
C 9	148	15.6	1013	17	US-10-264-237-205
C 10	115	12.1	557	11	US-09-969-034-4254
C 11	102	10.7	309	9	US-09-864-761-18145
C 12	98	10.3	483	9	US-09-864-761-1386
C 13	79	8.3	425	10	US-09-918-995-35878
C 14	75	7.9	376	9	US-09-983-965-1633
C 15	19	2.0	60	10	US-09-908-975-15332
C 16	9	0.9	201	18	US-10-719-993-10379
C 17	9	0.9	201	18	US-10-719-993-41266
C 18	9	0.9	201	19	US-10-741-600-27452
C 19	9	0.9	612	13	US-10-027-632-187795
C 20	9	0.9	612	17	US-10-027-632-187795
C 21	9	0.9	661	9	US-09-919-580-881
C 22	9	0.9	842	18	US-10-425-115-90494
C 23	9	0.9	1440	18	US-10-437-963-63191
C 24	9	0.9	1863	17	US-10-424-599-31853
C 25	9	0.9	4294	18	US-10-437-963-63187
C 26	9	0.9	6574	13	US-10-194-163-1097
C 27	9	0.9	9181	10	US-09-984-827-10
C 28	9	0.9	20284	17	US-10-452-510-21
C 29	9	0.9	20284	17	US-10-617-334-21
C 30	9	0.9	20284	18	US-10-744-465-21
C 31	9	0.9	20284	18	US-10-833-679-21
C 32	9	0.9	83719	18	US-10-388-838-45
C 33	9	0.9	98146	18	US-10-322-281-126
C 34	9	0.9	159138	18	US-10-719-993-6777
C 35	9	0.9	159138	19	US-10-741-600-17613
C 36	9	0.9	174481	19	US-10-662-613-6
C 37	9	0.9	183999	18	US-10-745-377-1
C 38	9	0.9	183999	18	US-10-872-113-1
C 39	9	0.9	303172	18	US-10-719-993-6890
C 40	8	0.8	25	19	US-10-719-900-49940
C 41	8	0.8	25	19	US-10-719-900-271965
C 42	8	0.8	25	19	US-10-719-900-520291
C 43	8	0.8	25	19	US-10-719-900-555961
C 44	8	0.8	25	19	US-10-719-900-737630
C 45	8	0.8	148	17	US-10-424-599-43421

ALIGNMENTS

RESULT 1  
US-10-921-707-25  
; Sequence 25, Application US/10921707  
; Publication No. US20050003447A1  
; GENERAL INFORMATION:  
; APPLICANT: INCYTE PHARMACEUTICALS, INC.  
; APPLICANT: LAL, Preeti  
; APPLICANT: TANG, Y. Tom  
; APPLICANT: YUE, Henry  
; APPLICANT: HILLMAN, Jennifer L.  
; APPLICANT: BANDMAN, Olga  
; APPLICANT: CORLEY, Neil C.  
; APPLICANT: GUEGLER, Karl J.  
; APPLICANT: PATTERSON, Chandra  
; APPLICANT: AZIMZAI, Yalda  
; APPLICANT: BAUGHN, Mariah R.  
; TITLE OF INVENTION: HUMAN CYTOSKELETON ASSOCIATED PROTEINS  
; FILE REFERENCE: PF-0594 PCT  
; CURRENT APPLICATION NUMBER: US/10/921,707  
; CURRENT FILING DATE: 2004-08-19

; PRIOR APPLICATION NUMBER: US/09/786,797  
 ; PRIOR FILING DATE: 2001-06-25  
 ; PRIOR APPLICATION NUMBER: 09/156,470; unassigned; 60/131,321  
 ; PRIOR FILING DATE: 1998-09-18; 1998-09-18; 1999-04-27  
 ; NUMBER OF SEQ ID NOS: 32  
 ; SOFTWARE: PERL Program  
 ; SEQ ID NO 25  
 ; LENGTH: 4527  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: misc\_feature  
 ; OTHER INFORMATION: Incyte ID No: 2363327  
 US-10-921-707-25

## Alignment Scores:

Pred. No.:	0	Length:	4527
Score:	711.00	Matches:	944
Percent Similarity:	98.44%	Conservative:	0
Best Local Similarity:	98.44%	Mismatches:	2
Query Match:	74.92%	Indels:	15
DB:	18	Gaps:	0

US-09-671-687A-3 (1-949) x US-10-921-707-25 (1-4527)

QY	4	GlyLeuTrpSerGlnGluLysValThrSerProTyrTrpGluGluArgIlePheTyrLeu	23
DB	270	GGCTTATGGAGCCCAAGAAAGTCACTTCACTCCCTACTGGGAAGAGCGGATTTTACTTG	329
QY	24	LeuLeuGlnGluCysSerValThrAspLysGlnThrGlnLysLeuLeuLysValProLys	43
DB	330	CTTCTTCAAGAAATGAGCGTTTACAGACAAACAAACACAAAGCTCTTAAAGTACCGAA-	388
QY	44	GlySerIleGlyGlnTyrIleGluAspArgSerValGlyHisSerArgIleProSerAla	63
DB	389	GGAGGTATAGACAGTATATTCAAGATCGTTCTGTGGGGCATTCAGGATTCCTTCGCA	448
QY	64	LysGlyLysLysAsnGlnIleGlyLeuLysIleLeuGluGlnProHisAlaValLeuPhe	83
DB	449	AAAGGCAAGAAATCAGATTGGATTAAATTTCTAGAGCAACCTCATGCGAGTCTCTTT	508
QY	84	ValAspGlu---AspValValGluIleAsnGluLysPheThrGluLeuLeuAlaIle	102
DB	509	GTTTATGAAAGGATGTTGTAGAGATAAATGAAAGTTTACAGAGTACTTTTGGCAATT	568
QY	103	ThrAsnCysGluGluArgPheSerLeuPheLysAsnArgAsnArgLeuSerLysGlyLeu	122
DB	569	ACCAATTTGTAGGAGAGGTTTCAAGCTGTGTTTAAAAACAGAAAAACAGACTAAGTAAAGGCCTC	628
QY	123	GlnIleAspValGlyCysProValLysValGlnLeuArgSerGlyGluGluLysPhePro	142
DB	629	CNAATACGTGGGCTGTCCTGTGAAGTACAGCTGAGATCTGGGGAAGAAAAATTTCT	688
QY	143	GlyValValArgPheArgGlyProLeuLeuAlaGluArgThrValSerGlyIlePhePhe	162
DB	689	GGAGTTGTACGCTTACAGAGACCCCTGTATACAGAGAGACAGTCTCCGGAATATTCTTT	748
QY	163	GlyValGlnLeuLeuGluGluArgGlyGlnGlyPheThrAspGlyValTyrGlnGly	182
DB	749	GGAGTTGAATTTGTGGAAGAGGTCGTGCTCAAGGTTTCACTGACGGGGTGTACCAAGGG	808
QY	183	LysGlnLeuPheGlnCysAspGluAspCysGly---PheValAlaLeuAspLysLeuGlu	201
DB	809	AAACAGCTTTTTCAGTGTGATGAAGATTGTGGCGTGTGTTTGTTCATTGGCAAGCTAGAA	868
QY	202	LeuIleGluAspAspThrAlaLeuGluSerAspTyrAlaGlyProGlyAspThrMet	221
DB	869	CTCATAGAGATGATGACACTGCAATTGGAAAGTGAATTACGAGGTCCTGGGACACAATG	928
QY	222	GlnValGluLeuProProLeuGluIleAsnSerArgValSerLeuLysGly-GlyGluTh	241
DB	929	CAGGTGCAACTCTCTCTTGGAAATAAATCCAGAGTTTCTTGAA--GGTTGGAGAAAC	987

QY	241	rIleGluSerGlyThrValIlePheCysAspValLeuProGlyLysGluSerLeuGlyTy	261
DB	988	AATAGAACTCGAAACAGTTATATTCTGTGATGTTTTTGCAGGAAAGAAAGCTTAGGATA	1047
QY	261	rPheValGlyValAspMetAspAsnProIleGlyAsnTrpAspGlyArgPheAspGlyVa	281
DB	1048	TTTTGTTGGTGGGACATGGATAACCTATTGGCAACTGGGATGGGAAGATTTGATGGAGT	1107
QY	281	1---LeuCysSerPheAlaCysValGluSerThrIleLeuLeuHisIleAsnAspIle1	300
DB	1108	GCAGCTTTGTAGTTTTTGGTGTGTTGAAGATCAATTTCTATTGCAATCATATATCAT	1167
QY	300	eProGluSerValThrGlnGluArgProProLysLeuAlaPheMetSerArgGlyVa	320
DB	1168	CCAGAGAGTGTGACGACAGGAAAGAGGCTCCCAAACTTGCCTTTATGTCAAGAGGTCT	1227
QY	320	lGlyAspLysGlySerSerSerHisAsnLysProLysAlaThrGlySerThrSerAspPr	340
DB	1228	TGGGACAAAGGTTTCATCCAGTCAATAAACCACAAAGGCTACAGGATCTACCTCAGACC	1287
QY	340	oGlyAsnArg---ArgSerGluLeuPheTyrThrLeuAsnGlySerSerValAspSerGl	359
DB	1288	TGGAATAAGAAACAGATCTGAATTTATTTATACCTTTAAATGGGTCTTCTGTTGACTACA	1347
QY	359	nProGlnSerLysSerLysAsnThrTrpTyrIleAspGluValAlaGluAspProAlaLys	379
DB	1348	ACCAATTCCAATCAAAAATACATGATGATGATGATGATGATGATGATGATGATGATGAT	1407
QY	379	sSerLeuThrGluIleSerThrAspPheAspArgSerSerProProLeuGlnProProPr	399
DB	1408	ATCTCTTACAGAGATATCTACAGACTTTGACCGTCTTTCACCAACCTCCAGCCTCTCC	1467
QY	399	oValAsnSerLeuThrThrGluAsnArgPheHisSerLeuProPheSerLeuThrLysMe	419
DB	1468	TGTGAACCTCCTGACCCAGAGAAACAGATTCCTTTTACCATTCAGTCTCACCAGAT	1527
QY	419	tProAsnThrAsnGlySerIleGlyHisSerProLeuSerLeuSerAlaGlnSerValMe	439
DB	1528	GCCCATACCAATGGAGATTTGGCCACAGTCCACTTTCTCTGTCAGCCAGCTGTGTAAT	1587
QY	439	tGluLeuLeuAsnThrAlaProValGlnGlnSerProProLeuAlaMetProProGlyAs	459
DB	1588	GGAGAGCTAAACACTGCACCCGCTCCAAGAGAGTCCACCTTGGCCATGCTCTCTGGGAA	1647
QY	459	nSerHisGlyLeuGluValGlySerLeuAlaGluValLysGluAsnProPheTyrGln	479
DB	1648	CTCATCATGCTTAGAAGTGGGCTCATTTGGCTGAAGTAAAGAGAACCTCTCTTCTATGG	1707
QY	479	yValIleArgTrpIleGlyGlnProProGlyLeuAsnGluValLeuAlaGlyLeuGluLe	499
DB	1708	GGTAATCCGTTGGATCGGTCCAGCCACAGGACTGAATGAAGTCTGCTGGAGCTGGAAC	1767
QY	499	uGluAspGluCysAlaGlyCysThrAspGlyThrPheArgGlyThrArgTyrPheThrCy	519
DB	1768	GGAGATGAGTGTGACGCTGTACGATGGAACCTTTCAGAGCACTCCGTAATTTCCACCTG	1827
QY	519	aAlaLeuLysLysAlaLeuPheValLysLeuLysSerCysArgProAspSerArgPheAl	539
DB	1828	TGCCCTGAAGAAGGCGCTGTTTGTGAAACTGAAAGAGTGCAGGCTGACTCTAGGTTTGC	1887
QY	539	aSerLeuGlnProValSerAsnGlnIleGluArgCysAsnSerLeuAlaPheGlyGlyTy	559
DB	1888	ATCATTTGAGCGCGTTTCCAAATCAGATTCAGCGCTGTAACCTTTTAGCATTTGGAGCTA	1947
QY	559	rLeuSerGluValValGluGluAsnThrProProLysMetGluLysGluGlyLeuGluI	579
DB	1948	CTTAAGTGAAGTAGTAGAAGAAATATCTCCACCAAAATGGAAAGAGGCTTGGAGAT	2007
QY	579	eMetIleGlyLysLysGlyIleGlnGlyHisTyrAsnSerCysTyrLeuAspSerTh	599
DB	2008	AATGATTGGGAAGAAAGGATCCAGGCTCATTAATTTCTGTTACTTAGACTCAAC	2067
QY	599	rLeuPheCysLeuPheAlaPheSerSerValLeuAspThrValLeuLeuArgProLysGl	619



```
Db 2068 CTTATTCTGCTATTGCTTTTGTAGTTCTGTTCTGGACACTGTTTACTTTAGACCCCAAGA 2127
Qy uLysAsnAspValGluTyrTyrSerGluThrGlnGluLeuLeuArgThrGluLeuValAs 639
Db 2128 AAAGAACGATGAGAAATATTAGTGAACCCAGAGCTACTGAGGACAGAAATGTTAA 2187
Qy nProLeuArgIleTyrGlyTyrValCysAlaThrLysIleMetLysLeuArgLysIleLe 659
Db 2188 TCCTCTGAGATATATGATATGTTGTGTCACAAAAATATGAAACTGAGGAAATACT 2247
Qy uGluLysValGluAlaIleSerGlyPheThrSerGluGluLysAspProGluGluPheLe 679
Db 2248 TCAAAAGGTGAGGCTGCATCAGGATTTACCTCTGAAGAAAAAGATCCTGAGGAATCTT 2307
Qy uAsnIleLeuPheHisIleLeuArgValGluProLeuLeuLysIleArgSerAlaG 699
Db 2308 GAATATCTGTTTCATCATATTTTAAAGGTAGAACCTTTGCTTAAAAATAGATCAGCAG 2367
Qy yGlnLysValGlnAspCysTyrPheTyrGlnIlePheMetGluLysAsnGluLysValG 719
Db 2368 TCAAAAGGTACAGATGTTGTTACTTCTATCAAAATTTTATGGAATAAATGAGAAATTGG 2427
Qy yValProThrIleGlnGlnLeuLeuGluTyrPheIleAsnSerAsnLeuLysPheAl 739
Db 2428 CGTCCCAACAATTCAGCAGTGTAGATGGTCTTTTATCAACAGTAACTCGAAATTCG 2487
Qy aGluAlaProSerCysLeuIleGlnMetProArgPheGlyLysAspPheLysLeuPh 759
Db 2488 AGAGGCACCATCATGCTGATTTATTCAGATGCTCGATTGGAAGAAGACTTTAAACTATT 2547
Qy eLysIysIlePheProSerLeuGluLeuAsnIleThrAspLeuLeuGluAspThrProAr 779
Db 2548 TAAAAAAATTTTCTCTCGAATTAATATAACAGATTTACTTTGAAGACACTCCCAAG 2607
Qy gGlnCysArgIleCysGlyGlyLeuAlaMetTyrGluCysArgGluCysTyrAspAspPr 799
Db 2608 ACAGTGGCGGATATGCGAGGCTTGCAATGATGATGATGATGATGATGATGATGATGAT 2667
Qy oAspIleSerAlaGlyLysIleLysGlnPheCysIysThrCysAsnThrGlnValHisLe 819
Db 2668 GGACATCTCAGCTGGAATAATCAAGCAGTTTGTAAAACTGCAACACTCAAGTCCACCT 2727
Qy uHisProLysArgLeuAsnHisLysTyrAsnProValSerLeuProLysAspLeuProAs 839
Db 2728 TCATCCGAAGAGGCTGAATCAATAATATAACCCAGTGTCACTTCCCAAGACTTACCCGA 2787
Qy pTyrAspTyrArgHisGlyCysIleProCysGlnAsnMetGluLeuPheAlaValLeuCy 859
Db 2788 CTGGACTGGAGACACGGCTGCATCCCTTCCAGATATGAGGATTTATTTGCTGTTCTCG 2847
Qy sIleGluThrSerHisTyrValAlaPheValLysTyrGlyLysAspAspSerAlaTrpLe 879
Db 2848 CATGAAACAAGCCACTATGTTGCTTTGTGAAGTATGGAAGGACGATTCTGCCTGGCT 2907
Qy uPhePheAspSerMetAlaAspArgAspGlyGlyGlnAsnGlyPheAsnIleProGlnVa 899
Db 2908 CTTCTTTGACAGCATGGCGGATCGGGATGGTGGTCAGAAATGGCTTCAACATTCCTCAAG 2967
Qy lThrProCysProGluValGlyGluTyrLeuLysMetSerLeuGluAspLeuHisSerLe 919
Db 2968 CACCCATGCCAGAAAGTAGGAGAGTACTTGAAGATGTTCTGGAAGACCTGATTCCTT 3027
Qy uAspSerArgArgIleGlnGlyCysAlaArgArgLeuLeuCysAspAlaTyrMetCysMe 939
Db 3028 GGACTCCAGGAGAAATCCAAGGCTGTGACAGAGACTGCTTTGTGATGATATATGTCAT 3087
Qy tTyrGlnSerProThrMetSerLeuTyrLys 949
Db 3088 GTACACAGAGTCCAAACAATGAGTTTGTACAAA 3118
```

RESULT 2

US-09-851-673-3

```
/ Sequence 3, Application US/09851673
/ Publication No. US20030165985A1
/ GENERAL INFORMATION:
/ APPLICANT: Derry, Jonathan
/ APPLICANT: Fanslow, William
/ APPLICANT: Dougall, William
/ TITLE OF INVENTION: SCREENING ASSAYS FOR AGONISTS OR ANTAGONISTS OF CD40 SIGNALING
/ FILE REFERENCE: 3198
/ CURRENT APPLICATION NUMBER: US/09/851,673
/ CURRENT FILING DATE: 2001-05-08
/ NUMBER OF SEQ ID NOS: 4
/ SOFTWARE: PatentIn version 3.0
/ SEQ ID NO 3
/ LENGTH: 5371
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (392)..(3262)
US-09-851-673-3
Alignment Scores:
Pred. No.: 0 Length: 5371
Score: 685.00 Matches: 948
Percent Similarity: 97.63% Conservative: 0
Best Local Similarity: 97.63% Mismatches: 1
Query Match: 72.18% Indels: 23
DB: 10 Gaps: 0
```

US-09-671-687A-3 (1-949) x US-09-851-673-3 (1-5371)

```
Qy 1 MetSerSerGlyLeuTyrSerGlnGluLysValThrSerProTyrTrpGluGluArgIle 20
Db 392 ATGAGTTCAAGCTTATGGAGCCCAAGAAAAAGTCATTCACCTTCTGGGAAGAGCGGATT 451
Qy 21 PheTyrLeuLeuLeuGlnGluCysSerValThrAspLysGlnThrGlnLysLeuLys 40
Db 452 TTTTACTTCTTCTTCAAGATGTCAGCGTTTACAGACAAACACAAAAAGCTCTTTAA 511
Qy 41 ValProLysGlySerIleGlyGlnTyrIleGlnAspArgSerValGlyHisSerArgIle 60
Db 512 GTACCCAGGGAAGTATAGGACAGTATATTCAAGATCGTTCTGTGGGGCATTTCAAGGATT 571
Qy 61 ProSerAlaLysGlyLysLysAsnGlnIleGlyLeuLysIleLeuGlnProHisAla 80
Db 572 CCTTCTGCAAAAGGCAAGAAAAATCAGATTGGATTAAAAATCTTAGAGCAACCTCATGCA 631
Qy 81 ValLeuPheValAspGlu--AspValValGluIleAsnGluLysPheThrGluLeuLeu 99
Db 632 GTTCTCTTTGTTGATGAAAGGATGTTGTAGAGATAAATGAAAAGTTTACAGAGTTACTT 691
Qy 100 LeuAlaIleThrAsnCysGluGluArgPheSerLeuPheLysAsnArgAsnArgLeuSer 119
Db 692 TTGGCAATTTACCAATTTGTGAGGAGAGGTTTCAGCTGTGTTAAAAACAGAACACAGACTAAGT 751
Qy 120 LysGlyLeuGlnIleAspValGlyCysProValLysValGlnLeuArgSerGlyGluGlu 139
Db 752 AAAGGCTCTCAAAATAGACGTGGGCTGCTCTGTGAAAGTACAGCTGAGATCTGGGGAAGAA 811
Qy 140 LysPheProGlyValValArgPheArgGlyProLeuLeuAlaGluArgThrValSerGly 159
Db 812 AAATTTCTGGAGTTGTACGCTTTCAGAGGACCCCTGTAGCAGAGGACAGACAGTCTCCGA 871
Qy 160 IlePhePheGlyValGluLeuLeuGluGlyArgGlyValGlnGlyPheThrAspGlyVal 179
Db 872 ATATCTTTGGAGTTGAATTTGCTGGAAGAGGTCGTGGTCAAGGTTTCACTGACGGGTG 931
Qy 180 TyrGlnGlyLysGlnLeuPheGlnCysAspGluAspCysGly---PheValAlaLeuAsp 198
Db 932 TACCAAGGGGAAACAGCTTTTTCAGTGTGATGAAGATTGTGGCGCTGTTGTGTGATGGAC 991
Qy 199 LysLeuGluLeuIleGluAspAspThrAlaLeuGluSerAspTyrAlaGlyProGly 218
```

Db 992 AAGCTAGAACCTCATAGAGATGATGACACTGCATTTGGAAAGTGATTAACGAGGCTCTGGG 1051  
Qy AspThrMetGlnValGluLeuProLeuGluLeuAenSerArgValSerLeuLysGly 238  
Db GACAAATGCAAGTCCGAATCTCTCTTTGGAAATAAACTCCAGAGTTCTTTGAA-GGT 1110  
Qy -GlyGluThrIleGluSerGlyThrValIlePheCysAspValLeuProGlyLysGluSe 258  
Db 1111 TGGAGAAACAATAGAAATCGAACAAGTTATATCTGTGATGTTTGGCCAGGAAAGAAAG 1170  
Qy rLeuGlyTyrPheValGlyValAspMetAspAsnProIleGlyAenTTPAspGlyArgPh 278  
Db 1171 CTTAGGATATTTTGTGTGTGGACATGATGAATACCTATTTGGCAACTGGGATGGAGATT 1230  
Qy eAspGlyVal---LeuCysSerPheAlaCysValGluSerThrIleLeuLeuHisIleAs 297  
Db 1231 TGATGGAGTGCAGCTTTTGTAGTTTTCGTGTGTGAAAGTACAAATCTATTGCACATCNA 1290  
Qy nAspIleIlePro-----GluSerValThrGlnGluArgArgProProLysLeuAl 314  
Db 1291 TGATATCATCCAGCTTTTATCAGAGAGTGTGACGACGAGGAAGAGGCGCTCCCAAACTTGC 1350  
Qy 314 aPheMetSerArgGlyValGlyAspLysGlySerSerSerHisAsnLysProLysAlaTh 334  
Db 1351 CTTTATGTCAAGAGGTGTGGGGACAAAGGTTTCATCCAGTCAATAATAACCAAGGCTAC 1410  
Qy 334 rGlySerThrSerAspProGlyAsnArg---ArgSerGluLeuPheTyrThrLeuAsnG1 353  
Db 1411 AGGATCTACCTCAGACCTCGGAAATAGAAACAGATCTGNAATTTATTTATACCTTAAATGG 1470  
Qy 353 ySerSerValAspSerGlnProGlnSerLysSerLysAsnThrThrTyrIleAspGluVa 373  
Db 1471 GTCTTCTGTGACTCACACACCAATCCAAATCAAAATAATACATGATGATGATGAGT 1530  
Qy 373 lAlaGluAspProAlaLysSerLeuThrGluLeSerThrAspPheAspArgSerSerPr 393  
Db 1531 TGCAGAACCCCTGCMAAATCTTTACAGAGATATCTACAGACTTTGACCGTCTTTCACC 1590  
Qy 393 oProLeuGlnProProValAsnSerLeuThrThrGluAsnArgPheHisSerLeuPr 413  
Db 1591 ACCACTCCAGCTCTCTCTGTGAATCTGACACCGGAGAACAGATTCACCTCTTTACC 1650  
Qy 413 oPheSerLeuThrLysMetProAsnThrAsnGlySerIleGlyHisSerProLeuSerIle 433  
Db 1651 ATTCACTCTCACCAGATGCCAATACCAATGGAAGTATTTGCCACAGTCCACTTTCCT 1710  
Qy 433 uSerAlaGlnSerValMetGluGluLeuAsnThrAlaProValGlnGluSerProProLe 453  
Db 1711 GTCAGCCAGTCTGTAATGGAAGAGCTAAACACTGCACCCGTCCAAGAGAGTCCACCCCT 1770  
Qy 453 uAlaMetProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLysGl 473  
Db 1771 GGCCATGCTCTCTGGGAATCACTAGTGTGTAAGTGGGCTCATTTGGCTGAAGTTAAGGA 1830  
Qy 473 uAsnProProPheTyrGlyValIleArgTrrIleGlyGlnProProGlyLeuAsnGluVa 493  
Db 1831 GAACCTCTCTTATATGGGTAAATCCGTGGATCGGTGACCGACACAGAGTGAATGAAGT 1890  
Qy 493 lLeuAlaGlyLeuGluLeuAspGluCysAlaGlyCysThrAspGlyThrPheArgG1 513  
Db 1891 GCTCGCTGGACTGGAACCTGGAAGATGAGTGTGAGGCTGTACGGATGGAACCTTCAGAGG 1950  
Qy 513 yThrArgTyrPheThrCysAlaLeuLysLysAlaLeuPheValLysLeuLysSerCysAr 533  
Db 1951 CACTCGGTATTTCACCTGTGCCCTGAAGAAGCGCTGTTTGTGAAACTGAAGAGCTGTCAG 2010  
Qy 533 gProAspSerArgPheAlaSerLeuGlnProValSerAsnGlnIleGluAtqCysAsnSe 553  
Db 2011 GCCTGACTCTAGTTTGCATCATTCAGCCGGTTTCCAAATCAGATTGAGCGCTGTAATCT 2070  
Qy 553 rLeuAlaPheGlyGlyTyrLeuSerGluValValGluGluAsnThrProProLysMetG1 573  
Db 2071 TTTAGCAITTTGGAGGCTACTTAAGTGAAGTAGTAGAAGAAATACTCCACCAAAATGGA 2130

Qy 573 ulysGluGlyLeuGluIleMetIleGlyLysLysGlyIleGlnGlyHisTyrAsnSe 593  
Db 2131 AAAAGAGGCTTGGAGATAATGATTGGGAAGAAAGAGCATCCAGGGTCAATACAAATTC 2190  
Qy 593 rCysTyrLeuAspSerThrLeuPheCysLeuPheAlaPheSerSerValLeuAspThrVa 613  
Db 2191 TTGTTACTTAGACTCAACCTTATTCTGCTTATTGCTTTAGTCTTGTCTGTGCACTGT 2250  
Qy 613 lLeuLeuArgProLysGluLysAsnAspValGluTyrTyrSerGluThrGlnGluLeuLe 633  
Db 2251 GTTACTTTAGACCCCAAGAAAGAACGATGTAGATATTATATAGTGAACCCCAAGAGCTACT 2310  
Qy 633 uArgThrGluIleValAsnProLeuArgIleTyrGlyTyrValCysAlaThrLysIleMe 653  
Db 2311 GAGGACAGAAATTTGTTAATCTCTCAGAAATATATGATATGTTGTGCCACAAAATTTAT 2370  
Qy 653 lLysLeuArgLysIleLeuGluLysValGluAlaSerGlyPheThrSerGluGluLys 673  
Db 2371 GAAACTGAGGAAAAATCTTGAAGGTGAGGCTGCATCAGGATTTACCTCTGAAGAAAA 2430  
Qy 673 sAspProGluGluPheLeuAsnIleLeuPheHisIleLeuArgValGluProLeuLe 693  
Db 2431 AGATCCTCAGGAAATCTTGAATATTCTGTTTCATCATATTTTAAAGGTAGAACCTTTGCT 2490  
Qy 693 ulysIleArgSerAlaGlyGlnLysValGlnAspCysTyrPheTyrGlnIlePheMetG1 713  
Db 2491 AAAAAATAGATCAGCAGGTCAAAAGGTACAGATTTGTTACTTCTATCAAAATTTTATGGA 2550  
Qy 713 ulysAsnGluLysValGlyValProThrIleGlnGlnLeuLeuGluTTPSerPheIleAs 733  
Db 2551 AAAAAATGAGAAAGTTGGCGTTCCCAAAATTCAGCAGTTGTTAGAAATGGTCTTTTATCAA 2610  
Qy 733 nSerAsnLeuLysPheAlaGluAlaProSerCysLeuIleIleGlnMetProArgPheG1 753  
Db 2611 CAGTAACCTGAAATTTGCAGAGGCCACCATCATGCTGATTTATTCAGATGCTCGATTTGG 2670  
Qy 753 yLysAspPheLysLeuPheLysLysIlePheProSerLeuGluLeuAsnIleThrAspLe 773  
Db 2671 AAAAACTTTAAACTATTTTAAAAAAATTTTCTCTCTCGGAATTTAAATATTAACAGATTT 2730  
Qy 773 uLeuGluAspThrProArgGlnCysArgIleCysGlyGlyLeuAlaMetTyrGluCysAr 793  
Db 2731 ACTTGAAGACACTCCACAGACAGTGCAGGATATGTGGAGGCTTGCAAATGATAGGTAG 2790  
Qy 793 gGluCysTyrAspAspProAspIleSerAlaGlyLysIleLysGlnPheCysLysThrCy 813  
Db 2791 AGAATGCTACGACGATCCGGACATCTCAGCTGGAATAATCAAGCAGTTTTGTAAAAACCTG 2850  
Qy 813 sAsnThrGlnValHisLeuHisProLysArgLeuAsnHisLysTyrAsnProValSerLe 833  
Db 2851 CAACACTCAAGTCCACCTTCATCCGAAGAGGTGAAATATAATAATAAACCCAGTGTCACT 2910  
Qy 833 uProLysAspLeuProAspTrrAspTrrArgHisGlyCysIleProCysGlnAsnMetG1 853  
Db 2911 TCCCAAGACTTACCCGACTGGGACTGGAGACACCGCTGCATCCCTTGCAGAAATATGA 2970  
Qy 853 uLeuPheAlaValLeuCysIleGluThrSerHisTyrValAlaPheValLysTyrGlyLy 873  
Db 2971 GTTATTTGCTGTCTCTGCATAGAAACAAGCCACTATGTTGCTTTGTAAGTATGGGAA 3030  
Qy 873 sAspAspSerAlaTrrPhePheAspSerMetAlaAspArgAspGlyGlyGlnAsnG1 893  
Db 3031 GGACGATTTGCTGGCTCTTCTTTGACAGCATGCGCGATCGGGATGGGTGAGTGTGAGAAATGG 3090  
Qy 893 yPheAsnIleProGlnValThrProCysProGluValGlyGluTyrLeuLysMetSerLe 913  
Db 3091 CTTCAACATTTCTCAAGTCACCCCATGCCAGAGTAGAGAGTACTTGAAGATGTCTCT 3150  
Qy 913 uGluAspLeuHisSerLeuAspSerArgArgIleGlnGlyCysAlaArgArgLeuLeuCy 933  
Db 3151 GGAAGACCTGCATTTCTTGGACTCCAGGAGAAATCCAGGCTGTGCACGAGACTGCTTTG 3210

Qy 933 sAspAlaTyrMetCysMetTyrGlnSerProThrMetSerLeuTyrLys 949  
Db 3211 TGATGCATATATGTCATGTACACAGTCCAAACATGAGTTGTACAAA 3259

## RESULT 3

US-10-755-889-489  
; Sequence 489, Application US/10755889  
; Publication No. US20040171823A1  
; GENERAL INFORMATION:  
; APPLICANT: Bristol-Myers Squibb Company  
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES ASSOCIATED WITH THE NF-KB  
; TITLE OF INVENTION: PATHWAY  
; FILE REFERENCE: D0284 NP  
; CURRENT APPLICATION NUMBER: US/10/755,889  
; CURRENT FILING DATE: 2004-01-13  
; PRIOR APPLICATION NUMBER: U.S. 60/440,068  
; PRIOR FILING DATE: 2003-01-14  
; PRIOR APPLICATION NUMBER: U.S. 60/469,757  
; PRIOR FILING DATE: 2003-05-12  
; NUMBER OF SEQ ID NOS: 823  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 489  
; LENGTH: 5371  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-755-889-489

Alignment Scores:  
Pred. No.: 0 Length: 5371  
Score: 685.00 Matches: 948  
Percent Similarity: 97.63% Conservative: 0  
Best Local Similarity: 97.63% Mismatches: 1  
Query Match: 72.18% Indels: 23  
DB: 18 Gaps: 0

US-09-671-687A-3 (1-949) x US-10-755-889-489 (1-5371)

Qy 1 MetSerSerGlyLeuTyrSerGlnGluLysValThrSerProTyrTrpGluArgIle 20  
Db 392 ATGAGTTCAGGCTTATGGAGCCAAAGAAAGTCACTTCACCTACTTGGGAAGCGGATT 451  
Qy 21 PheTyrLeuLeuGlnGluCysSerValThrAspLysGlnThrGlnLysLeuLeuLys 40  
Db 452 TTTTACTTCTCTTCAAGATGCGAGTTTACAGACAAACAAACAAAGTCTCTTAAA 511  
Qy 41 ValProLysGlySerIleGlnTyrIleGlnAspArgSerValGlyHisSerArgIle 60  
Db 512 GTACCGAAGGGAAGTATAGGACAGTATATTCAAGATCGTTCTGTGGGCAATCAAGATT 571  
Qy 61 ProSerAlaLysGlyLysLysGlnIleGlyLeuLysIleLeuGlnProHisAla 80  
Db 572 CCTTCGCAAGGCAAGAAATCAGATTGGATTAAAAATTTCTAGAGCAACCTCATGCA 631  
Qy 81 ValLeuPheValAspGlu---AspValValGluIleAsnGluLysPheThrGluLeuLeu 99  
Db 632 GTTCTCTTGTGTGAAGAGATGTTGTAGAGATAAATGAAGATTCTACAGATTACTT 691  
Qy 100 LeuAlaIleThrAsnCysGluArgPheSerLeuPheLysAsnArgAsnArgLeuSer 119  
Db 692 TTGGCAATTACCAATTGTGAGGAGAGTTACGCTGTTTAAAAACAGAAACAGACTAAGT 751  
Qy 120 LysGlyLeuGlnIleAspValGlyCysProValLysValGlnLeuArgSerGlyGluGlu 139  
Db 752 AAAGCCCTCCAAATAGAGCTGGGCTGCTCCTGTGAAAGTACAGCTGAGATCTGGGAAGAA 811  
Qy 140 LysPheProGlyValValArgPheArgGlyProLeuLeuAlaGluArgThrValSerGly 159  
Db 812 AAATTTCTGGAGTTGTACGCTTACAGAGACCCCTGTTAGCAGAGAGGACAGTCTCCGGA 871  
Qy 160 IlePhePheGlyValGluLeuLeuGluGluArgGlyArgGlyGlnGlyPheThrAspGlyVal 179  
Db 872 ATATTCTTGGAGTTGAATTGCTGGAGAGAGTCTGTGTCAGGTTTCACTGACGGGGTG 931

Qy 180 TyrGlnGlyLysGlnLeuPheGlnCysAspGluAspCysGly---PheValAlaLeuAsp 198  
Db 932 TACCAAGGGAACACAGCTTTTTCAGTGTGATGAAGATTGTGCGCTGTTGTGTGATTGGAC 991  
Qy 199 LysLeuGluLeuIleGluAspAspThrAlaLeuGluSerAspTyrAlaGlyProGly 218  
Db 992 AAGCTAGAATCATAGAAGATGATGACATGCAATTTGGAAGATGATTACGAGGTCCTGGG 1051  
Qy 219 AspThrMetGlnValGluLeuProLeuGluIleAsnSerArgValSerLeuLysGly 238  
Db 1052 GACACAATGAGGTGCAATCTCTCTTTGGAATAAATCCAGAGTTCTTTGAA--GGT 1110  
Qy 239 -GlyGluThrIleGluSerGlyThrValIlePheCysAspValLeuProGlyLysGluSe 258  
Db 1111 TGGAGAAACAATAGAAATCTGGAACAGTTATATCTGTGATGTTTGGCAGAAAGAAAG 1170  
Qy 258 rLeuGlyTyrPheValGlyValAspMetAspAsnProIleGlyAsnTrpAspGlyArgPh 278  
Db 1171 CTTAGGATATTTTGTGTGTGACATGGAATACCTTATTGGCAACTGGGATGGAAGATT 1230  
Qy 278 eAspGlyVal---LeuCysSerPheAlaCysValGluSerThrIleLeuLeuHisIleAs 297  
Db 1231 TGATGAGTGCAGCTTTGTAGTTTTCGCTGTGTGAAAGTACAAATCTATTGCACATCAA 1290  
Qy 297 nAspIleIlePro-----GluSerValThrGlnGluArgArgProProLysLeuAl 314  
Db 1291 TGNATATCATCCAGCTTTTATCAGAGAGTGTGACGAGAGAAAGGAGGCTCCCAACTTGC 1350  
Qy 314 aPheMetSerArgGlyValGlyAspLysGlySerSerSerHisAsnLysProLysAlaIth 334  
Db 1351 CTTTATGTCAAGAGGTGTTGGGACAAAGGTTTCATCCAGTCAATAATAAACCAAGGCTAC 1410  
Qy 334 rGlySerThrSerAspProGlyAsnArg---ArgSerGluLeuPheTyrThrLeuAsnG 353  
Db 1411 AGGATCTACCTCAGACCCCTGGAATAGAAACAGATCTGAATATTTTATACCTTAAATGG 1470  
Qy 353 ySerSerValAspSerGlnProGlnSerLysSerLysAsnThrTyrTyrIleAspGluVa 373  
Db 1471 GTCTTCTGTGTGACTCAACACCACCAATCCAAATCAAAAATATACATGTTGATGTAAGT 1530  
Qy 373 lAlaGluAspProAlaLysSerLeuThrGluIleSerThrAspPheAspArgSerSerPr 393  
Db 1531 TGCAGAGAGCCCTGCAAAATCTTTACAGAGATATCTACAGACTTTTGACCGTCTTCCACC 1590  
Qy 393 oProLeuGlnProProProValAsnSerLeuThrGluAsnArgPheHisSerLeuPr 413  
Db 1591 ACCACTCCAGCTCTCTCTGTGAACCTCACTGACCCAGAGAACAGATTCACACTCTTTACC 1650  
Qy 413 oPheSerLeuThrLysMetProAsnThrAsnGlySerIleGlyHisSerProLeuSerLe 433  
Db 1651 ATTCAGTCTCACCAAGATGCCCAATACCAATGGAAGTATTGGCCACAGTCCACTTCTCT 1710  
Qy 433 uSerAlaGlnSerValMetGluLeuAsnThrAlaProValGlnGluSerProProLe 453  
Db 1711 GTACGCCCACTCTGTATATGAGAGAGTAAACACTGACCCGCTCCAGAGAGTCCACCTT 1770  
Qy 453 uAlaMetProProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLysG 473  
Db 1771 GGCCATGCTCTCTGGGAACCTCACATGCTCTAGAAGTGGGCTCATTTGCTGAAGTAAAGGA 1830  
Qy 473 uAsnProProPheTyrGlyValIleArgTrpIleGlyGlnProProGlyLeuAsnGluVa 493  
Db 1831 GAACCTCTCTTCTATGGGGTAAATCCGTTGGATCGGTCAGCCACAGGACTGAATGAAGT 1890  
Qy 493 lLeuAlaGlyLeuGluLeuAspGluCysAlaGlyCysThrAspGlyThrPheArgG 513  
Db 1891 GTCGCTGAGCTGGAACTCGAAGATGAGTGTGAGGCTGTACGATGGAACTTCAGAGG 1950  
Qy 513 yThrArgTyrPheThrCysAlaLeuLysLysAlaLeuPheValLysLeuLysSerCysAr 533  
Db 1951 CACTCGGTATTTTCACTGTGCTGGAAGAGGCGCTGTTTGTGAAACTGAAGAGCTGCAG 2010  
Qy 533 gProAspSerArgPheAlaSerLeuGlnProValSerAsnGlnIleGluArgCysAsnSe 553

```
Db 2011 GCCTGACTCTAGGTTTGCAATCAGCGGTTTCCCAATCAGATTGAGCGCTGTAATCTC 2070
Qy rLeuAlaPheGlyGlyTyrLeuSerGluValValGluGluAsnThrProProLysMetG1 573
Db 2071 TTTAGCATTTGGAGGCTACTTAAGTAGAGTAGTAGAGAAATATCTCCACCAAAATGGA 2130
Qy uLysGluGlyLeuGluLeuMetIleGlyLysLysGlyLysGlyLysGlyLysGlyLys 593
Db 2131 AAAAGAGGCTTGGAGATTAATGATTGGGAAGAAAGGCATCCAGGGTCATTACAAATTC 2190
Qy rCysTyrLeuAspSerThrLeuPheCysLeuPheAlaPheSerSerValLeuAspThrVa 613
Db 2191 TTGTACTTAGACCAACCTTATTCTGCTTATTCTGCTTTTGTAGTTCTGTCGACACTGT 2250
Qy lLeuLeuArgProLysGluLysAsnAspValGluTyrTyrSerGluThrGlnGluLeuLe 633
Db 2251 GTTACTTAGACCAACAAAGAAAGACAGTAGTAGAATATTATAGTGAACCCAGAGACTACT 2310
Qy uArgThrGluIleValAsnProLeuArgIleTyrGlyTyrValCysAlaThrLysIleMe 653
Db 2311 GAGGACAGAAATTTGTAATCTCTCAGAAATATATGATATGTTGTGTCACCAAAATTTAT 2370
Qy tLysLeuArgLysIleLeuGluLysValGluAlaAlaSerGlyPheThrSerGluGluLy 673
Db 2371 GAAACTGAGGAAATACTTTGAAAGAGTGGAGGCTGCATCAGGATTTACCTCTGAAGAAA 2430
Qy sAspProGluGluPheLeuAsnIleLeuPheHisIleLeuArgValGluProLeuLe 693
Db 2431 AGATCCTCAGGAAATTTCTGGAATATTCTGTTCATCATATTTTAAAGGGTAGAACCTTGT 2490
Qy uLysIleArgSerAlaGlyGlnLysValGlnAspCysTyrPheTyrGlnIlePheMetG1 713
Db 2491 AAAATAAGATCAGAGCTCAAAAGGTACAGATTGTTACTTCTATCAAAATTTTATGGA 2550
Qy uLysAsnGluLysValGlyValProThrIleGlnGlnLeuLeuGluTyrSerPheIleAs 733
Db 2551 AAAAAATGAGAAAGTTGGCGTTCACCAATTCAGCAGTGTGTAGAAATGGTCTTTTATCAA 2610
Qy nSerAsnLeuLysPheAlaGluAlaProSerCysIleuIleGlnMetProArgPheG1 753
Db 2611 CAGTAACCTGAAATTTGAGAGGACCAATCATGCTGTGATTATTCAGATGCCCTCGATTTGG 2670
Qy yLysAspPheLysLeuPheLysLysIlePheProSerLeuGluLeuAsnIleThrAspLe 773
Db 2671 AAAAGACTTTAACTATTATTAATAAATTTTCTCTCTGGAATTAATATACAGATT 2730
Qy uLeuGluAspThrProArgGlnCysArgIleCysGlyGlyLeuAlaMetTyrGluCysAr 793
Db 2731 ACTTGAAGACACTCCACAGACAGTCCGGATATGTGAGGGCTTGCATATGATGAGTGTAG 2790
Qy gGluCysTyrAspAspProAspIleSerAlaGlyLysIleLysGlnPheCysLysThrCy 813
Db 2791 AGAATGCTACGACGATCCGACATCTCAGCTGGAATAATCAAGCAGTTTGTAAACCTG 2850
Qy sAsnThrGlnValHisLeuHisProLysArgLeuAsnHisLysTyrAsnProValSerle 833
Db 2851 CAACACTCAAGTCCACCTTCATCCGAAGAGGCTGAATCATAAATATAACCCAGTGTCACT 2910
Qy uProLysAspLeuProAspTyrAspTrpArgHisGlyCysIleProCysGlnAsnMetG1 853
Db 2911 TCCCAAGACTTACCCGACTGGGACTGGAGACACCGCTGCATCTCTGCCAGAAATATGA 2970
Qy uLeuPheAlaValLeuCysIleGluThrSerHisTyrValAlaPheValLysTyrGlyLy 873
Db 2971 GTTATTTGCTGTTCTCTGCATAGAAACAAGCCACTATGTTGCTTTTGTGAAGTATGGAA 3030
Qy sAspAspSerAlaTrpLeuPhePheAspSerMetAlaAspArgAspGlyGlyGlnAsnG1 893
Db 3031 GGACGATTTCTCCTGGCTCTTCTTTTGACAGCATGGCCGATCGGATGGTGGTGCAGATGG 3090
Qy yPheAsnIleProGlnValThrProCysProGluValGlyGluTyrLeuLysMetSerle 913
```

```
Db 3091 CTTCAACATTCTCAAGTCACCCCATGCCAGAAAGTAGGAGAGTACTTGAAGATGTCTCT 3150
Qy uGluAspLeuHisSerLeuAspSerArgArgIleGlnGlyCysAlaArgArgLeuLeuCy 933
Db 3151 GGAAGACCTGCATCTCTTGACTCCAGAGAGATCCAGGCTGTGCACGAGACTGCTTGG 3210
Qy sAspAlaTyrMetCysMetTyrGlnSerProThrMetSerLeuTyrLys 949
Db 3211 TGATCATATATGTGATGTACAGAGTCCAACATGAGTTTGTACAAA 3259

RESULT 4
US-10-037-270-290
; Sequence 290, Application US/10037270
; Publication No. US20030104529A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yunding
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: Tillinghast, John
; APPLICANT: Dmanac, Radoje T.
; TITLE OF INVENTION: No. US20030104529A1el Nucleic Acids and
; FILE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/10/037,270
; CURRENT FILING DATE: 2002-01-04
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1104
; SOFTWARE: pt FL_genes Version 1.0
; SEQ ID NO 290
; LENGTH: 2523
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (81)..(2138)
US-10-037-270-290

Alignment Scores:
Pred. No.: 0 Length: 2523
Score: 641.00 Matches: 668
Percent Similarity: 99.55% Conservative: 0
Best Local Similarity: 99.55% Mismatches: 0
Query Match: 67.54% Indels: 3
DB: 15 Gaps: 0

US-09-671-687A-3 (1-949) x US-10-037-270-290 (1-2523)
Qy 282 LeuCysSerPheAlaCysValGluSerThrIleLeuLeuHisIleAsnAspIlelePro 301
Db 129 CTTTCTAGTTTTCGCTGTGTGAAAGTAGTACAATTCTATTGCATCATCAATGATATCATCCA 188
Qy 302 GluSerValThrGlnGluArgProProLysLeuAlaPheMetSerArgGlyValGly 321
Db 189 GAGAGTGTGACGACGAGAAAGGAGGCTCCCAAACTTGCCTTTATGTCAAGAGGTGTGG 248
Qy 322 AspLysGlySerSerHisAsnLysProLysAlaThrGlySerThrSerAspProGly 341
Db 249 GACAAGGTTTATCCAGTCATTAATAACCAAGAGGTACAGGATCTACCTCAGACCTGGA 308
```



; NUMBER OF SEQ ID NOS: 1104  
 ; SOFTWARE: pt\_FL\_genes Version 1.0  
 ; SEQ ID NO 290  
 ; LENGTH: 2523  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (81)..(2138)  
 US-10-117-722-290

Alignment Scores:  
 Pred. No.: 0 Length: 2523  
 Score: 641.00 Matches: 668  
 Percent Similarity: 99.55% Conservativity: 0  
 Best Local Similarity: 99.55% Mismatches: 0  
 Query Match: 67.54% Indels: 3  
 DB: 17 Gaps: 0

US-09-671-687A-3 (1-949) x US-10-117-722-290 (1-2523)

Qy 282 LeuCySerPheAlaCysValGluSerThrIleLeuLeuHisIleAsnAspIlePro 301  
 Db 129 CTTTGTAGTTTGGTGTGGAAAGTACAAATCTATTGCAATCAATGATATCATCCA 188  
 Qy 302 GluSerValThrGlnGluArgProProLysLeuAlaPheMetSerArgGlyValGly 321  
 Db 189 GAGAGTGTCGCGAGGAAAGGAGGCTCCCAACTTGGCTTTATGTCAAGAGGTCTGGG 248  
 Qy 322 AspLysGlySerSerSerHisAsnLysProLysAlaThrGlySerThrSerAspProGly 341  
 Db 249 GACAAAGGTTTCATCCAGTCATAATAAACCAAGAGGTACAGGATCTACCTCAGACCCCTGA 308  
 Qy 342 AsnArg---ArgSerGluLeuPheTyThrLeuAsnGlySerSerValAspSerGlnPro 360  
 Db 309 AATGAAACACATCTGAATATTTTATACCTTAAATGGGTCTCTGTGTGACTCACAACCA 368  
 Qy 361 GlnSerLysSerLysAsnThrTptTyIleAspGluValAlaGluAspProAlaLysSer 380  
 Db 369 CAATCCAAATCAAAAATACATGGTACATGATGATGGTGGAGAGACCTCGCAAAATCT 428  
 Qy 381 LeuThrGluIleSerThrAspPheAspArgSerSerProProLeuGlnProProVal 400  
 Db 429 CTTACAGAGATATCTACAGACTTTGACCGTTCTTCAACCACCACTCCAGCCTCCTCTGTG 488  
 Qy 401 AsnSerLeuThrThrGluAsnArgPheHisSerLeuProPheSerLeuThrLysMetPro 420  
 Db 489 AACTCACTGACCCGAGAACAGATTCACCTCTTTACCATTCCAGTCTCACCAGAGATGCC 548  
 Qy 421 AsnThrAsnGlySerIleGlyHisSerProLeuSerLeuSerAlaGlnSerValMetGlu 440  
 Db 549 AATACCANTGGAAGTATTGGCCACAGTCACCTTTCTGTGACGCCAGTCTGTATATGGA 608  
 Qy 441 GluLeuAsnThrAlaProValGlnGluSerProProLeuAlaMetProProGlyAsnSer 460  
 Db 609 GAGCTAAACACTGCACCCGTCACAGAGAGTCCACCTTGGCCATCGCTCTGGGAACCTCA 668  
 Qy 461 HisGlyLeuGluValGlySerLeuAlaGluValLysGluAsnProProPheTyThrVal 480  
 Db 669 CATGGTCTAGAAAGTGGGCTCATTTGGCTGAAAGTTAAGGAGAACCTCTCTTCTATGGGTA 728  
 Qy 481 IleArgTptIleGlyGlnProProGlyLeuAsnGluValLeuAlaGlyLeuLeuGlu 500  
 Db 729 ATCCGTTGGATCGGTGACGCCACAGACTGAATGAAGTGTCTGCTGGGACTGGAACTGAA 788  
 Qy 501 AspGluCysAlaGlyCysThrAspGlyThrPheArgGlyThrArgTyThrPheThrCysAla 520  
 Db 789 GATGAGTGTGAGGCTGTACGGATGGAACCTTCAGAGGCATCTCGGTATTTTCACTTGCC 848  
 Qy 521 LeuLysLysAlaLeuPheValLysLeuLysSerCysArgProAspSerArgPheAlaSer 540  
 Db 849 CTGAAGAAGGCGCTGTTTGTGAAACTGAAGAGCTGCAGGCTGCTCTAGGTTTGCATCA 908

Qy 541 LeuGlnProValSerAsnGlnIleGluArgCysAsnSerSerLeuAlaPheGlyGlyTyLeu 560  
 Db 909 TTGCAGCCGGTTTCCAATCAGATTGAGCGCTGTAACCTCTTTAGCATTTGGAGGCTACTTA 968  
 Qy 561 SerGluValValGluGluAsnThrProProLysMetGluLysGluGlyLeuGluLeuMet 580  
 Db 969 AGTGAAGTAGTAGAAGAAATACTCCACCAAAATGGAAGAGGCTTGGAGATAATG 1028  
 Qy 581 IleGlyLysLysLysGlyIleGlnGlyHisTyAsnSerCysTyThrLeuAspSerThrLeu 600  
 Db 1029 ATTGGGAAGAGAAGGCACTCAGGCTCATTAACAATCTTGTACTTAGACTCAACCTTA 1088  
 Qy 601 PheCysLeuPheAlaPheSerSerValLeuAspThrValLeuLeuArgProLysGluLys 620  
 Db 1089 TTCTGCTTATTGCTTTTAGTTCTGTTCTGGACACTGTGTTACTTAGACCCCAAGAAAG 1148  
 Qy 621 AsnAspValGluTyTyThrSerGluThrGlnGluLeuLeuArgThrGluIleValAsnPro 640  
 Db 1149 AACGATGTAGAATATTATAGTGAAACCCCAAGAGCTACTGAGGACAGAAATTTGTTAATCCT 1208  
 Qy 641 LeuArgIleTyThrGlyTyValCysAlaThrLysIleMetLysLeuArgLysIleLeuGlu 660  
 Db 1209 CTGAGATATATGATATGATATGTTGTCACAAAATATGAAACTGAGGAAATATCTTGAA 1268  
 Qy 661 LysValGluAlaAlaSerGlyPheThrSerGluGluLysAspProGluGluPheLeuAsn 680  
 Db 1269 AAGGTGGAGGCTGCATCAGGATTTTACCTCTGAAGAAAAGATCTCTGAGGATTTCTTGAAT 1328  
 Qy 681 IleLeuPheHisIleLeuArgValGluProLeuLeuLysIleArgSerAlaGlyGln 700  
 Db 1329 ATTCTGTTTCATCATATTTTAAGGGTGAACCTTTGCTGAAAAAATTAAGATCAGCAGTCAA 1388  
 Qy 701 LysValGlnAspCysTyThrPheTyGlnIlePheMetGluLysAsnGluLysValGlyVal 720  
 Db 1389 AAGGTACAGGATTTGTTACTTCTATCAAAATTTTATGGAAAAAATGAGAAAGTTGCGCTT 1448  
 Qy 721 ProThrIleGlnGlnLeuLeuGluTrpSerPheIleAsnSerAsnLeuLysPheAlaGlu 740  
 Db 1449 CCCCAATTCAGCAGTGTGTAGATGGTCTTTTATCAACAGTAACCTGAAATTTGCAGAG 1508  
 Qy 741 AlaProSerCysLeuIleGlnMetProArgPheGlyLysAspPheLysLeuPheLys 760  
 Db 1509 GCACCATCATGCTGATTATTACAGATGCTCGATTTGGAAAGACTTTTAAACTATTTAAA 1568  
 Qy 761 LysIlePheProSerLeuGluLeuAsnIleThrAspLeuLeuGluAspThrProArgGln 780  
 Db 1569 AAAATTTTCTCTCTCGAAATTAATAATAACAGATTTTACTTGAAGACACTCCACAGAC 1628  
 Qy 781 CysArgIleCysGlyGlyLeuAlaMetTyThrGluCysArgGluCysTyThrAspAspProAsp 800  
 Db 1629 TGCCGANTATGTGGAGGCTTGCATGATGATGATGATGATGATGATGATGATGATGATGAT 1688  
 Qy 801 IleSerAlaGlyLysIleLysGlnPheCysLysThrCysAsnThrGlnValHisLeuHis 820  
 Db 1689 ATCTCAGCTGGAAAAATCAAGCAGTTTTTGTAAAACTCTCAACACTCAAGTCCACCTTCAT 1748  
 Qy 821 ProLysArgLeuAsnHisLysTyThrAsnProValSerLeuProLysAspLeuProAspTrp 840  
 Db 1749 CCGAAGAGGCTGAATCATATAATAATAAACCCAGTGTCACTTCCCAAGACTTACCCGACTGG 1808  
 Qy 841 AspTrpArgHisGlyCysIleProCysGlnAsnMetGluLeuPheAlaValLeuCysIle 860  
 Db 1809 GACTGGAACACGCGCTGCATCCCTTGGCAGAAATATGGAGTTATTTGCTGTCTCTGCTATA 1868  
 Qy 861 GluThrSerHisTyValAlaPheValLysTyThrLysAspAspSerAlaTrpLeuPhe 880  
 Db 1869 GAAACAGCCACTATGTTGCTTTGTGAAGTATGGGAAGGAGCATCTTCCCTCGCTCTCTTC 1928  
 Qy 881 PheAspSerMetAlaAspArgAspGlyGlyGlnAsnGlyPheAsnIleProGlnValThr 900  
 Db 1929 TTTGACAGCATGGCCGATCGGATGGTGGTGGATGGCTTCAACATTTCTCAAGTCACTACC 1988  
 Qy 901 ProCysProGluValGlyGluTyThrLysMetSerLeuGluAspLeuHisSerLeuAsp 920



Db 1989 CCATGCCAGAAGTAGGAGACTTGAAGATGCTCTGGAAGACCTGCATTCCTTGGAC 2048  
Qy 921 SerArgArgileGlnGlyCysAlaArgArgLeuLeuCysAspAlaTyrMetCysMetTyr 940  
Db 2049 TCCAGAGAGATCCAGGCTGTGCACGAGACCTGTTTGATGCATATATGTCATGTAC 2108  
Qy 941 GlnSerProThrMetSerLeuTyrLys 949  
Db 2109 CAGAGTCCAAACAATGAGTTTGTACAAA 2135

RESULT 6

US-10-264-237-204/c  
; Sequence 204, Application US/10264237  
; Publication No. US20040009491A1  
; GENERAL INFORMATION:  
; APPLICANT: Birse et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

; FILE REFERENCE: PA31PI  
; CURRENT APPLICATION NUMBER: US/10/264,237  
; CURRENT FILING DATE: 2002-10-04  
; PRIOR APPLICATION NUMBER: PCT/US01/16450  
; PRIOR FILING DATE: 2001-05-18  
; PRIOR APPLICATION NUMBER: US 60/205,515  
; PRIOR FILING DATE: 2000-05-19  
; NUMBER OF SEQ ID NOS: 2876  
; SOFTWARE: PatentIn Ver. 3.1  
; SEQ ID NO 204  
; LENGTH: 1151

; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (15)..(15)

; OTHER INFORMATION: n equals a,t,g, or c  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1092)..(1092)

; OTHER INFORMATION: n equals a,t,g, or c  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1113)..(1113)

; OTHER INFORMATION: n equals a,t,g, or c  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1136)..(1136)

; OTHER INFORMATION: n equals a,t,g, or c  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1142)..(1142)

; OTHER INFORMATION: n equals a,t,g, or c  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1142)..(1142)

; OTHER INFORMATION: n equals a,t,g, or c  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1142)..(1142)

; OTHER INFORMATION: n equals a,t,g, or c  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1142)..(1142)

; OTHER INFORMATION: n equals a,t,g, or c  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1142)..(1142)

; OTHER INFORMATION: n equals a,t,g, or c  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1142)..(1142)

; OTHER INFORMATION: n equals a,t,g, or c  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1142)..(1142)

; OTHER INFORMATION: n equals a,t,g, or c  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1142)..(1142)

; OTHER INFORMATION: n equals a,t,g, or c  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1142)..(1142)

; OTHER INFORMATION: n equals a,t,g, or c  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1142)..(1142)

; OTHER INFORMATION: n equals a,t,g, or c  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1142)..(1142)

; OTHER INFORMATION: n equals a,t,g, or c  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1142)..(1142)

; OTHER INFORMATION: n equals a,t,g, or c  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1142)..(1142)

; OTHER INFORMATION: n equals a,t,g, or c  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1142)..(1142)

; OTHER INFORMATION: n equals a,t,g, or c  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1142)..(1142)

Db 876 GATGTAGAATATATATAGTGAACCCAGAGACTACTGAGGACAGAAATTTGTAATCCTCTG 817  
Qy 642 ArgileTyrGlyTyrValCysAlaThrLysIleMetLysLeuArgLysIleLeuGluLys 661  
Db 816 AGAATATATGATATGTGTGTGCACAAAATATGAACTGAGGAAAATACCTTGAAAG 757  
Qy 662 ValGluAlaAlaSerGlyPheThrSerGluGluLysAspProGluGluPheLeuAsnIle 681  
Db 756 GTGAGGCTGCATCAGGATTTACCTCTCTGAAGAAAAAGATCCTGAGGAATTTCTTGAATATT 697  
Qy 682 LeuPheHisIleLeuArgValGluProLeuLeuLysIleArgSerAlaGlyGlnLys 701  
Db 696 CTGTTTCATCATATTTTAAGGGTAGAACCTTTGCTAAAAATAGATCAGCAGGTCAAAG 637  
Qy 702 ValGlnAspCysTyrPheTyrGlnIlePheMetGluLysAsnGluLysValGlyValPro 721  
Db 636 GTACAAGATTGTTACTTCTTCTATCAANTTTTATGGAAAAAATGAGAAGTTGGCGTCCC 577  
Qy 722 ThrIleGlnGlnLeuLeuGluTyrSerPheIleAsnSerAsnLeuLysPheAlaGluAla 741  
Db 576 ACAATTGAGGAGTTGTTAGATGCTTTTATCAACAGTAACTGAAATTTGCAGAGGCA 517  
Qy 742 ProSerCysLeuIleGlnMetProArgPheGlyLysAspPheLysLeuPheLysLys 761  
Db 516 CCATCATGCTGATTTTACAGATGCTCGATTTTGGAAAAAGACTTTAAACTATTAAAAA 457  
Qy 762 IlePheProSerLeuGluLeuAsnIleThrAspLeuLeuGluAspThrProArgGlnCys 781  
Db 456 ATTTTTCCTTCTCTGGAATTAATAATTAACAGATTTTCTTGAAGACATCCAGACAGTGC 397  
Qy 782 ArgileCysGlyGlyLeuAlaMetTyrGluCysArgGluCysTyrAspAspProAspIle 801  
Db 396 CGGATATGTGGAGGCTTGCAATGTATGAGTGTAGAGAACTGCTACGAYGATCCGACATC 337  
Qy 802 SerAlaGlyLysIleLysGlnPheCysLysThrCysAsnThrGlnValHisLeuHisPro 821  
Db 336 TCAGCTGGAAAAATCAAGCAGTTTGTGTAACCTGCAACACTCAAGTCCACCTTCATCCG 277  
Qy 822 LysArgLeuAsnHisLysTyrAsnProValSerLeuProLysAspLeuProAspTyrAsp 841  
Db 276 AAGAGGCTGAATCATTAATAATTAACCCAGTGTCACTTCCCAAGACTTTACCCGACTGGAC 217  
Qy 842 TrpArgHisGlyCysIleProCysGlnAsnMetGluLeuPheAlaValLeuCysIle 860  
Db 216 TGGAGACACGGCTGCATCCCTTCCCAAGATATGAGGATTTTGTCTGTCTCTGCATA 160

RESULT 7

US-09-969-034-4253/c  
; Sequence 4253, Application US/09969034  
; Publication No. US2004011068A1  
; GENERAL INFORMATION:  
; APPLICANT: Burgess, Christopher C.  
; APPLICANT: Astle, Jon H.  
; APPLICANT: Carroll, Eddie III  
; APPLICANT: Catino, Theodore J.  
; APPLICANT: Dwivedi, Poornima  
; APPLICANT: Molino, Gary A.  
; APPLICANT: Thiagalingam, Arunthathi  
; APPLICANT: Lewis, Marcia E.  
; TITLE OF INVENTION: Nucleic Acid Sequences Differentially  
; FILE REFERENCE: 1657/1032  
; CURRENT APPLICATION NUMBER: US/09/969,034  
; CURRENT FILING DATE: 2001-10-02  
; PRIOR APPLICATION NUMBER: 60/237,271  
; PRIOR FILING DATE: 2000-02-10  
; NUMBER OF SEQ ID NOS: 4494  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4253  
; LENGTH: 617  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:



```
; NAME/KEY: misc_feature
; LOCATION: 517, 586
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (691)..(691)
; OTHER INFORMATION: n is unknown.
US-10-761-370-3

Alignment Scores:
Pred. No.: 1.18e-163 Length: 2116
Score: 165.00 Matches: 178
Percent Similarity: 99.44% Conservative: 0
Best Local Similarity: 99.44% Mismatches: 0
Query Match: 17.39% Indels: 1
DB: 16 Gaps: 0

US-09-671-687A-3 (1-949) x US-10-761-370-3 (1-2116).
QY 387 AspPheAspArgSerSerProLeuGlnProProValAsnSerLeuThrThrGlu 406
DB 15 GACTTTGACCGTCTCTTCCACCACCTCCAGCCCTCTCTGTGAACTCACTGACCCGAG 74
QY 407 AsnArgPheHisSerLeuProPheSerLeuThrLysMetProAsnThrAsnGlySerIle 426
DB 75 AACAGATTCCACTCTTTACCATTCAGTCTCACCAGATGCCAATACCAATGGAGTATT 134
QY 427 GlyHisSerProLeuSerLeuSerAlaGlnSerValMetGluGluLeuAsnThrAlaPro 446
DB 135 GSCCAGACGTCCACTTTCTCTGTCCAGCCAGTCTGTAATGGAAGAGCTAAACACTGCACCC 194
QY 447 ValGlnGluSerProProLeuAlaMetProProGlyAsnSerHisGlyLeuGluValGly 466
DB 195 GTCCAAGAGAGTCCACCCCTTGCCCATGCTCTCTGGAACTCATAGTGGTCTAGAGTGGGC 254
QY 467 SerLeuAlaGluValLysGluAsnProPheTyrGlyValIleArgTyrIleGlyGln 486
DB 255 TCATTGGCTGAAGTTAAGAGAGACCTCTCTTCTATGGGGTAAATCCGTGGATCCGTGAG 314
QY 487 ProProGlyLeuAsnGluValLeuAlaGlyLeuGluLeuGluAspGluCysAlaGlyCys 506
DB 315 CCACAGGACTGAATGAAGTGTCTGCTGGACTGGAACTGGAAGATGAGTGTGACGCTGT 374
QY 507 ThrAspGlyThrPheArgGlyThrArgTyrPheThrCysAlaLeuLysLysAlaLeuPhe 526
DB 375 ACGGATGGAACCTTCAGAGGCACCTCGGTATTTCACCTGTGCCCTGAAGAAGCGCTGTTT 434
QY 527 VallysLeuLysSerCysArgProAspSerArgPheAlaSerLeuGlnProValSerAen 546
DB 435 GTGAACTGAAGAGCTGCAGGCTGACCTAGTTCATTCATTCAGCCGGTTCCTCAAT 494
QY 547 Gln-IleGluArgCysAsnSerLeuAlaPheGlyTyrLeuSerGluValVal 564
DB 495 CAAGATTGACGCGCTTAACCTCTTAGCATTTGGAGGCTACTTAAGTGAAGTAGTG 549

RESULT 9
US-10-264-237-205
; Sequence 205, Application US/10264237
; Publication No. US20040009491A1
; GENERAL INFORMATION:
; APPLICANT: Birse et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA131P1
; CURRENT APPLICATION NUMBER: US/10/264,237
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/16450
; PRIOR FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: US 60/205,515
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 2876
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 205

; NAME/KEY: misc_feature
; LOCATION: 517, 586
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (691)..(691)
; OTHER INFORMATION: n is unknown.
US-10-761-370-3

Alignment Scores:
Pred. No.: 1.67e-170 Length: 617
Score: 171.00 Matches: 171
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 18.02% Indels: 0
DB: 11 Gaps: 0

US-09-671-687A-3 (1-949) x US-09-969-034-4253 (1-617)
QY 737 LysPheAlaGluAlaProSerCysLeuIleIleGlnMetProArgPheGlyLysAspPhe 756
DB 514 AAATTTCAGAGGACCACTATGCTGATTTATTCAGATGCCCTCGAATTTGGAAAGACTTT 455
QY 757 LysLeuPheLysLysIlePheProSerLeuGluLeuAsnIleThrAspLeuLeuGluAsp 776
DB 454 AAACATTTTAAAAAATTTTCTCTCTGGAATTAATATAACAGATTACTTGAAGAC 395
QY 777 ThrProArgGlnCysArgIleCysGlyGlyLeuAlaMetTyrGluCysArgGluCysTyr 796
DB 394 ACTCCAGACAGTGGCGATATGTGGAGGCTTGCATATGATGAGTAGAGATGCTTAC 335
QY 797 AspAspProAspIleSerAlaGlyLysIleLysGlnPheCysLysThrCysAsnThrGln 816
DB 334 GACCATCCGGACATCTCAGCTGGAAAAATCAAGCAGTTTGTGAAACCTGCACACTCAA 275
QY 817 ValHisLeuHisProLysArgLeuAsnHisLysTyrAsnProValSerLeuProLysAsp 836
DB 274 GTCCACCTTCATCCGAAGAGCTGATCATTAATATACCCAGTGTCTACTTCCCAAGAC 215
QY 837 LeuProAspTyrAspTyrArgHisGlyCysIleProCysGlnAsnMetGluLeuPheAla 856
DB 214 TTACCCGACTGGGACTGGAGACACGCTGCATCCCTTGCAGAAATATGGAGTTATTGCT 155
QY 857 ValLeuCysIleGluThrSerHisTyrValAlaPheValLysTyrGlyLysAspSer 876
DB 154 GTTCTCTGATAGAAACAGCCATCATGTTGCTTTGTGAAGTAGTGGAGAGGACATTCT 95
QY 877 AlaTyrLeuPhePheAspSerMetAlaAspArgAspGlyGlyGlnAsnGlyPheAsnIle 896
DB 94 GCCTGGCTCTCTTTGACAGCATGCCCATCGGATGGTGGTGCAGATGGCTTCAACATT 35
QY 897 ProGlnValThrProCysProGluValGlyGlu 907
DB 34 CCTCAAGTCACCCCATGCTCCAGAGTAGGAGAG 2

RESULT 8
US-10-761-370-3
; Sequence 3, Application US/10761370
; Publication No. US20040219615A1
; GENERAL INFORMATION:
; APPLICANT: WALLACH, David
; APPLICANT: KOVALENKO, Andrei
; TITLE OF INVENTION: MODULATORS OF THE FUNCTION OF RECEPTORS OF THE TNF/NGF RECEPTOR F
; FILE REFERENCE: WALLACH=27
; CURRENT APPLICATION NUMBER: US/10/761,370
; CURRENT FILING DATE: 2004-01-22
; PRIOR APPLICATION NUMBER: US/09/646,403
; PRIOR FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: IL 123758
; PRIOR FILING DATE: 1998-03-19
; PRIOR APPLICATION NUMBER: PCT/IL99/00158
; PRIOR FILING DATE: 1999-03-18
; PRIOR APPLICATION NUMBER: IL 126024
; PRIOR FILING DATE: 1998-09-01
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.1
```

```
; LENGTH: 1013
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (5)..(5)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (9)..(9)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (307)..(307)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (948)..(948)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (984)..(984)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1004)..(1004)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-264-237-205

Alignment Scores:
Pred. No.:      5,85e-146      Length:      1013
Score:          148.00      Matches:      284
Percent Similarity: 98.61%      Conservative: 0
Best Local Similarity: 98.61%      Mismatches: 2
Query Match:      15.60%      Indels:      4
DB:              17          Gaps:      0

US-09-671-687A-3 (1-949) x US-10-264-237-205 (1-1013)

Qy 358 SerGlnProGlnSerLysSerLysAsnThrTyrTyrIleAspGluValAlaGluAspPro 377
Db 12 TCACAACCAACATCCAAATCAAAAAATACATGGTACATTCATGATGAAGTTGCAGAACCCCT 71
Qy 378 AlaLysSerLeuThrGluLeuSerThrAspPheAspArgSerSerProLeuGlnPro 397
Db 72 GCAAAATCTCTACAGAGATATCTACAGACTTTGACCGTCTTCCACCACCTCCAGCCT 131
Qy 398 ProProValAsnSerLeuThrThrGluAsnArgPheHisSerLeuProPheSerLeuThr 417
Db 132 CCTCTGTGAACCTACTGACCAACCGAGACAGATTCACCTCTTTACCATTCAGTCTCACC 191
Qy 418 LysMetProAsnThrAsnGlySerIleGlyHisSerProLeuSerLeuSerAlaGlnSer 437
Db 192 AAGATGCCCAATACCAATGGAAGTATTGGCCACAGTCCACTTCTCTGTGAGCCAGTCT 251
Qy 438 ValMetGluGluLeuAsnThrAlaProValGlnGlnSerProProLeuAlaMetPro-Pr 457
Db 252 GTAATGGAAGAGCTAAACACTGCACCCGTCGAAGAGAGTCCACCTTGGCCATGCN-TCC 310
Qy 457 oGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGlyValLysGluAsnProProPh 477
Db 311 TGGGACATCATGGTCTAGAAAGTGGGCTCATGTGCTGAAGTAAAGGAAACCCCTCTTT 370
Qy 477 eTyrGlyValIleArgTrrIleGlyGlnProProGlyLeuAsnGluValLeuAlaGlyLe 497
Db 371 CTATGGGGTAATCCGTTGGATCGGTGAGCCAGGAGTGAATGAAGTCTCGCTGGACT 430
Qy 497 uGluLeuGluAspGluCysAlaGlyCysThrAspGlyThrPheArgGlyThrArgTyrPh 517
Db 431 GGAACCTGGAAGATGAGTGTGAGGCTGTACGGATGGAACCTTTCAGAGGACCTCGGTATT 490
Qy 517 eThrCysAlaLeuLysLysAlaLeuPheValLysLeuLysSerCysArgProAspSerAr 537
```

```
Db 491 CACCTGTGCCCTGAAGAARGCGCTGTTGTGAAAACTGAAGAGCTGCAGGCGCTGACTCTAG 550
Qy 537 gPheAlaSerLeuGlnProValSerAsnGlnIleGluArgCysAsnSerLeuAlaPheGl 557
Db 551 GTTTGCATCATTCGACGCCGTTTCCATCAGATTGAGCGCTGTAACTCTTTAGCATTTGG 610
Qy 557 yGlyTyrLeuSerGluValValGluGluAsnThrProTyrLeuMetGluLysGluGlyLe 577
Db 611 AGGCTACTTAAGTGAAGTAGTAGAAGAAAAATCTCCACCAAAATGCAAAAAGAGGCTT 670
Qy 577 uGluIleMetIleGlyLysLysGlyIleGlnGlyHisTyrAsnSerCysTyrLeuAs 597
Db 671 GGAGATAATGATTGGGAAGAAGAGCATCCAGGCTCATTAATCTTTGTACTTTAGA 730
Qy 597 pSerThrLeuPheCysLeuPheAla-PheSerSerValLeuAspThrValLeuLeuArgP 617
Db 731 CTCACCTTATTCCTGCTTATTTSC-TTTTGTCTGTCTTGACACTGTGTTACTTTAGAC 789
Qy 617 rLysGluLysAsnAspValGluTyrTyrSerGluThrGlnGluLeuLeuArgThrGluI 637
Db 790 CCAAAGAAAAGAACGATGTAGATATATATAGTGAACCCCAAGAGCTACTGAGGACAGAA 849
Qy 637 leValAsnProLeuArgIle 643
Db 850 TTGTTAATCTCTGAGAAIT 869

RESULT 10
US-09-969-034-4254/c
; Sequence 4254, Application US/09969034
; Publication No. US20040110668A1
; GENERAL INFORMATION:
; APPLICANT: Burgees, Christopher C.
; APPLICANT: Astle, Jon H.
; APPLICANT: Carroli, Eddie III
; APPLICANT: Catino, Theodore J.
; APPLICANT: Dwivedi, Poornima
; APPLICANT: Molino, Gary A.
; APPLICANT: Thiagalingam, Arunthathi
; APPLICANT: Lewis, Marcia E.
; TITLE OF INVENTION: Nucleic Acid Sequences Differentially
; TITLE OF INVENTION: Expressed in Cancer Tissue
; FILE REFERENCE: 1657/1032
; CURRENT APPLICATION NUMBER: US/09/969,034
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: 60/237,271
; PRIOR FILING DATE: 2000-02-10
; NUMBER OF SEQ ID NOS: 4494
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4254
; LENGTH: 557
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 429, 468, 523
; OTHER INFORMATION: n = A,T,C or G
US-09-969-034-4254

Alignment Scores:
Pred. No.:      2,85e-111      Length:      557
Score:          115.00      Matches:      141
Percent Similarity: 98.60%      Conservative: 0
Best Local Similarity: 98.60%      Mismatches: 0
Query Match:      12.12%      Indels:      2
DB:              11          Gaps:      0

US-09-671-687A-3 (1-949) x US-09-969-034-4254 (1-557)

Qy 767 GluLeuAsnIleThrAspLeuLeuGluAspThrProArgGlnCysArgIleCysGly 786
Db 426 GAATTAATATAACAGATTTACTTGAAGACACTCCACAGACAGTCCCGGATATGTGAGGG 367
Qy 787 Leu-AlaMetTyrGluCysArgGluCysTyrAsp-AspProAspIleSerAlaGlyIleI 806
```

Db 366 CTTTGCATGATGAGTGTAGAGATGTTACCGCATCCGACATCTCAGCTGAAAAA 307  
Qy leLysGlnPheCysValThrCysAlaThrGlnValHisLeuHisProLysArgLeuAsnH 826  
Db 306 TCAAGCAGTTTGTAAACCTGCAACACTCAAGTCCACTTCATCCGAGAGGCTGAATC 247  
Qy isLysTyrAsnProValSerLeuProLysAspLeuProAspTyrAspTyrArgHisGlyC 846  
Db 246 ATAATATAACCCAGTGTCACTTCCCAAGACTTACCGACTGGAGCTGGAGACAGGCT 187  
Qy ysIleProCysGlnAsnMetGluLeuPheAlaValLeuCysIleGluThrSerHisTyrV 866  
Db 186 GCATCCCTTGCAGAAATAGGAGTATTTGCTGTTCTCTGCATAGAAACAAGCCACTATG 127  
Qy alAlaPheValLysTyrGlyLysAspAspSerAlaTyrPheLeuPheAspSerMetAlaA 886  
Db 126 TTGCTTTTGTGAAGTATGGAGGACGATTCTGCCTGCTCTCTTTTACAGCATGGCGG 67  
Qy sPaArgAspGlyGlyGlnAsnGlyPheAsnIleProGlnValThrProCysProGluValG 906  
Db 66 ATCGGAGTGTGTCAGATGGCTTCAACATTCCTCAAGTACCCCATGCCGAGAGTAG 7  
Qy 906 lyGlu 907  
Db 6 GAGAG 2

## RESULT 11

US-09-864-761-18145/c  
; Sequence 18145, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharon G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Chen, Wensheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; FILE REFERENCE: Aeonica-X-1  
; CURRENT APPLICATION NUMBER: US/09/864,761  
; CURRENT FILING DATE: 2001-05-23  
; PRIOR APPLICATION NUMBER: US 60/180,312  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: US 60/207,456  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/632,366  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: GB 24263.6  
; PRIOR FILING DATE: 2000-10-04  
; PRIOR APPLICATION NUMBER: US 60/236,359  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: PCT/US01/00666  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00667  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00664  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00669  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00665  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00668  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00663  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00662  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00661  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00670  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: US 60/234,687  
; PRIOR FILING DATE: 2000-09-21

; PRIOR APPLICATION NUMBER: US 09/608,408  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: US 09/774,203  
; PRIOR FILING DATE: 2001-01-29  
; NUMBER OF SEQ ID NOS: 49117  
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 18145  
; LENGTH: 309  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURES:  
; OTHER INFORMATION: MAP TO AC007728.1  
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 0.98  
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.92  
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.5  
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.3  
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1  
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 0.89  
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1  
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.3  
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.4  
; OTHER INFORMATION: EST HUMAN HIT: A1130924.1, EVALUE 0.00e+00  
; OTHER INFORMATION: SWISSPROT HIT: Q10427, EVALUE 1.70e-01  
; OTHER INFORMATION: NT HIT: g111430965, EVALUE 0.00e+00  
US-09-864-761-18145

## Alignment Scores:

Pred. No.:	9.51e-98	Length:	309
Score:	102.00	Matches:	102
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	10.75%	Indels:	0
DB:	9	Gaps:	0

US-09-671-687A-3 (1-949) x US-09-864-761-18145 (1-309)

Qy 387 AspPheAspArgSerProProLeuGlnProProValAsnSerLeuThrThrGlu 406  
Db 308 GACTTTGACCGTTCTTCCACCACCTCCAGCTCTCTCTGTGAATCTACTGACCACCGAG 249  
Qy 407 AsnArgPheHisSerLeuProPheSerLeuThrLysMetProAsnThrAsnGlySerIle 426  
Db 248 AACGATTCCACTCTTTACCATTCAGTCTCACCAGATGCCCAATACCAATGGAAGTATT 189  
Qy 427 GlyHisSerProLeuSerLeuSerAlaGlnSerValMetGluGluLeuAsnThrAlaPro 446  
Db 188 GGCACAGTCCACTTTCTCTGTCAGCCAGTCTGTAAATGGAAGAGCTAAACACTGCACCC 129  
Qy 447 ValGlnGluSerProProLeuAlaMetProProGlyAsnSerHisGlyLeuGluValGly 466  
Db 128 GTCCAAAGAGAGTCCACCTTTGGCCATGCTCTCTGGAACTCACAATGGTCTAGAAGTGGGC 69  
Qy 467 SerLeuAlaGluValLysGluAsnProProPheTyrGlyValIleArgTyrIleGlyGln 486  
Db 68 TCATTGGCTGAAGTTAAGAGAACCTCTCTTCTATGGGTAATCCGTTGGATCGGTGAG 9  
Qy 487 ProPro 488  
Db 8 CCACCA 3

## RESULT 12

US-09-864-761-1386/c  
; Sequence 1386, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharon G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Chen, Wensheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; FILE REFERENCE: Aeonica-X-1  
; CURRENT APPLICATION NUMBER: US/09/864,761

```

; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 1386
; LENGTH: 483
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC007728.1
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 0.98
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.92
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 0.89
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.4
US-09-864-761-1386

Alignment Scores:
Pred. No.: 2,47e-93 Length: 483
Score: 98.00 Matches: 98
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 10.33% Indels: 0
DB: 9 Gaps: 0

US-09-671-687A-3 (1-949) x US-09-864-761-1386 (1-483)

Qy 402 SerLeuThrGluAsnArgPheHisSerLeuProPheSerLeuThrIlyMetProAn 421
Db 483 TCATGACCCGAGAAACAGATTCACCTCTTTACCATTGAGTCTCACCAGATGCCCAAT 424
Qy 422 ThrAsnGlySerIleGlyHisSerProLeuSerLeuAlaGlnSerValMetGluGlu 441
Db 423 ACCAATGGAAGTATTGGCCACAGTCCACTTTCTGTGACCCAGTCTGTAATGGAAGAG 364
```

```

Qy 442 LeuAsnThrAlaProValGlnGluSerProProLeuAlaMetProProGlyAsnSerHis 461
Db 363 CTAAACACTGCACCCGCTCAAGAGAGTCCACCTTGGCCATGCTCTCGGAAGTACCAT 304
Qy 462 GlyLeuGluValGlySerLeuAlaGluValLysGluAsnProProPheTyrGlyValIle 481
Db 303 GGTCTAGAGTGGGCTCATTTGGCTGAAGTTAAGGAGAACCTCTCTTCTATGGGGTAATC 244
Qy 482 ArgTTrpIleGlyGlnProProGlyLeuAsnGluValLeuAlaGlyLeuGluLeu 499
Db 243 CGTTGGATCGTGCAGCCACGAGACTGAATGAAGTCTGCTGAGTGGAACTG 190
```

## RESULT 13

```

US-09-918-995-35878
; Sequence 35878, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; TITLE OF INVENTION: FROM VARIOUS CDNA LIBRARIES
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 35878
; LENGTH: 425
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(425)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-35878
```

## Alignment Scores:

```

Pred. No.: 2,83e-73 Length: 425
Score: 79.00 Matches: 79
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 8.32% Indels: 0
DB: 10 Gaps: 0
```

US-09-671-687A-3 (1-949) x US-09-918-995-35878 (1-425)

```

Qy 565 GluGluAsnThrProProLysMetGluLysGluGlyLeuGluIleMetIleGlyIlyLys 584
Db 3 GAAGAAATATCTCCACCAAAATGGAAGAGGCTTGAGATAATGATTGGGAAGAAG 62
Qy 585 LysGlyIleGlnGlyHisTyrAsnSerCysTyrLeuAspSerThrLeuPheCysLeuPhe 604
Db 63 AAAGGCAATCCAGGTCATTACAAATCTTGTTACTTAGACTCAACCTATTCTGCTATTT 122
Qy 605 AlaPheSerSerValLeuAspThrValLeuLeuArgProGlyGluLysAsnAspValGlu 624
Db 123 GCTTTTAGTTCTGTTCTGGACACTGTGTTACTTAGACCCAAAGAAAGACGATGTAGAA 182
Qy 625 TyrTyrSerGluThrGlnGluLeuLeuArgThrGluIleValAsnProLeuArgIle 643
Db 183 TATTATAGTAAACCCAGAGCTACTGAGCAGAAATTTGTTAATCCTCTGGAATTT 239
```

## RESULT 14

```

US-09-983-965-1633
; Sequence 1633, Application US/09983965
; Patent No. US20020137160A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
```

;; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND  
;; FILE OF INVENTION: MUSCLE AND FAT DEPOSITION  
;; FILE REFERENCE: 37-21(10297)C  
;; CURRENT APPLICATION NUMBER: US/09/983,965  
;; CURRENT FILING DATE: 2001-10-26  
;; PRIOR APPLICATION NUMBER: US 09/465,231  
;; PRIOR FILING DATE: 1999-12-15  
;; PRIOR APPLICATION NUMBER: US 60/113,678  
;; PRIOR FILING DATE: 1998-12-17  
;; NUMBER OF SEQ ID NOS: 5912  
;; SEQ ID NO 1633  
;; LENGTH: 376  
;; TYPE: DNA  
;; ORGANISM: Bos taurus  
;; FEATURE:  
;; OTHER INFORMATION: Clone ID: 47-LIB2809-020-Q1-E1-D4  
US-09-983-965-1633

## Alignment Scores:

Pred. No.:	4.33e-69	Length:	376
Score:	75.00	Matches:	75
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	7.90%	Indels:	0
DB:	9	Gaps:	0

US-09-671-687A-3 (1-949) x US-09-983-965-1633 (1-376)

Qy	730	SerPheIleAnSerAsnLeuLysPheAlaGluAlaProSerCysLeuIleIleGlnMet	749
Db	86	TCCTTTATCAACAGTAGTAACCTGCTGAGGCACCATCATGCTGATTATTTCAGATG	145
Qy	750	ProArgPheGlyLysAspPheLysLeuPheLysLysIlePheProSerLeuGluLeuAen	769
Db	146	CCTGATTGGAAAGACTTTAAACTCTTTAAAAAAATTTTCCTCTCTGGAACATAAT	205
Qy	770	IleThrAspLeuLeuGluAspThrProArgGlnCysArgIleCysGlyGlyLeuAlaMet	789
Db	206	ATAACAGATTTACTTGNAGACACCCCGAGCAGTCCGGATATGTGGAGGGCTTGCATG	265
Qy	790	TyrGluCysArgGluCysTyrAspAspProAspIleSerAlaGly	804
Db	266	TATGAGTGTAGAGATGCTATGATGACCTGCATCTCAGCTGGC	310

## RESULT 15

US-09-908-975-15332  
; Sequence 15332, Application US/09908975  
; Publication No. US20030165843A1  
; GENERAL INFORMATION:  
; APPLICANT: SHOSHAN, Avi  
; APPLICANT: WASSERMAN, Alon  
; APPLICANT: MINTZ, Eli  
; APPLICANT: MINTZ, Lia  
; APPLICANT: FAIGLER, Simchon  
; TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPLICE  
; FILE REFERENCE: 36688-0005  
; CURRENT APPLICATION NUMBER: US/09/908,975  
; CURRENT FILING DATE: 2001-07-20  
; PRIOR APPLICATION NUMBER: US 60/287,724  
; PRIOR FILING DATE: 2001-05-02  
; PRIOR APPLICATION NUMBER: US 60/221,607  
; PRIOR FILING DATE: 2000-07-28  
; NUMBER OF SEQ ID NOS: 32337  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 15332  
; LENGTH: 60  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-908-975-15332

## Alignment Scores:

Pred. No.:	1.48e-10	Length:	60
------------	----------	---------	----

Score:	19.00	Matches:	19
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	2.00%	Indels:	0
DB:	10	Gaps:	0

US-09-671-687A-3 (1-949) x US-09-908-975-15332 (1-60)

Qy 534 ProAspSerArgPheAlaSerLeuGlnProValSerAsnGlnIleGluArgCysAsn 552  
Db 2 CCTGACTCTAGTTTGCAATCATTCGACGGGTTTCCATTCAGATTCAGGCGCTGTAAAC 58

Search completed: April 17, 2005, 14:12:47  
Job time : 1193 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: April 17, 2005, 01:25:02 ; Search time 1076 Seconds  
(without alignments)  
5221.036 Million cell updates/sec

Title: US-09-671-687A-3

Perfect score: 949

Sequence: 1 MSSGLWSQKVTSPYWEERI.....RLLCDAYMCMYQPTMSLYK 949

Scoring table:

OLIGO  
Xgapop 60.0 , Xgapext 60.0  
Ygapop 60.0 , Ygapext 60.0  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 4390206 seqs, 2959870667 residues

Word size: 1

Total number of hits satisfying chosen parameters: 8776198

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame+\_p2n.model -DEV=xl  
-Q=/cgn2\_1/USPTO\_spool\_p/US09671687/runat\_15042005\_163037\_11942/app\_query.fasta\_1.1095  
-DB=N Geneseq\_16Dec04 -QFMT=fastap -SUFFIX=Oligo.rng -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09671687@cgn\_1\_1\_708@runat\_15042005\_163037\_11942 -NCPU=6 -ICPU=3  
-NO MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : N Geneseq\_16Dec04:\*

- 1: Geneseqn1980s:\*
- 2: Geneseqn1990s:\*
- 3: Geneseqn2000a:\*
- 4: Geneseqn2001a:\*
- 5: Geneseqn2001b:\*
- 6: Geneseqn2002a:\*
- 7: Geneseqn2002b:\*
- 8: Geneseqn2003a:\*
- 9: Geneseqn2003b:\*
- 10: Geneseqn2003c:\*
- 11: Geneseqn2003d:\*
- 12: Geneseqn2004a:\*
- 13: Geneseqn2004b:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	754	79.5	3302	12	Adq95917 T cell ac
2	754	79.5	6831	10	Adc24898 Human bre
3	711	74.9	4527	3	Aaa08589 Human cyt
4	685	72.2	5371	10	Abv75394 Human CYL
5	685	72.2	5371	13	Adrl4488 Human NF-

6	667	70.3	3302	12	ADQ95915	Adq95915 T cell ac
7	641	67.5	2523	4	AAI58410	AAI58410 Human pol
8	641	67.5	2523	5	ADQ98620	ADQ98620 DNA encod
9	641	67.5	2523	9	ADB48380	Novel hum
10	641	67.5	4716	4	AAI1617	CDNA enco
11	608	64.1	2845	4	AAH18625	AAH18625 Human CDN
12	598	63.0	3311	12	ADQ95919	Adq95919 T cell ac
13	418	44.0	2569	4	AAH18478	AAH18478 Human CDN
14	402	42.4	4286	4	AAI160196	AAI160196 Human pol
15	279	29.4	1151	6	ABL89642	ABL89642 Human pol
16	262	27.6	2488	4	AAI1081	CDNA enco
17	171	18.0	617	6	ABQ60558	ABQ60558 Human col
18	165	17.4	2116	2	AAZ07515	AAZ07515 Human RAP
19	148	15.6	1013	6	ABL89643	ABL89643 Human pol
20	124	13.1	837	4	AAH07749	AAH07749 Human CDN
21	123	13.0	758	4	AAH05084	AAH05084 Human CDN
22	115	12.1	557	6	ABQ60559	ABQ60559 Human col
23	106	11.2	318	12	ADK71948	Adk71948 Human ori
24	102	10.7	309	4	AAI20692	AAI20692 Probe #10
25	102	10.7	309	4	ABA65741	ABA65741 Human foe
26	102	10.7	309	4	AAI45907	AAI45907 Probe #14
27	102	10.7	309	4	ABA47848	ABA47848 Human bre
28	102	10.7	309	4	ABA32825	ABA32825 Probe #11
29	102	10.7	309	4	AAK39883	AAK39883 Human bon
30	102	10.7	309	4	AAK14143	AAK14143 Human bra
31	102	10.7	309	5	AAI06390	AAI06390 Probe #63
32	102	10.7	309	6	ABS13981	ABS13981 Human gen
33	101	10.6	303	12	ADK71946	Adk71946 Human ori
34	98	10.3	483	4	AAI11479	AAI11479 Probe #14
35	98	10.3	483	4	ABA53149	ABA53149 Human foe
36	98	10.3	483	4	AAI32754	AAI32754 Probe #14
37	98	10.3	483	4	ABA42721	ABA42721 Human bre
38	98	10.3	483	4	ABA22920	ABA22920 Probe #13
39	98	10.3	483	4	AAK26850	AAK26850 Human bon
40	98	10.3	483	4	AAK01396	AAK01396 Human bra
41	98	10.3	483	5	AAI01393	AAI01393 Probe #13
42	98	10.3	483	6	ABS01449	ABS01449 Human gen
43	96	10.1	288	12	ADK71950	Adk71950 Human ori
44	91	9.6	273	12	ADK71944	Adk71944 Human ori
45	86	9.1	1160	5	AAS90524	DNA encod

ALIGNMENTS

RESULT 1  
ADQ95917  
ID ADQ95917 standard; cDNA; 3302 BP.  
XX  
AC ADQ95917;  
XX  
DT 07-OCT-2004 (first entry)  
XX  
DE T cell activation associated cDNA #48.

ss; gene; anti-allergic; antiarthritic; antiasthmatic; antidiabetic;  
anti-HIV; antimicrobial; antirheumatic; immunosuppressive;  
neuroprotective; gene therapy; T cell activation; diagnosis;  
autoimmune disease; rheumatoid arthritis; asthma; multiple sclerosis;  
diabetes; allergic disease; infectious disease; AIDS; chronic rejection;  
organ; bone-marrow transplant.

OS Homo sapiens.

Key Location/Qualifiers  
CDS 243..3104  
/\*tag= a

WO2004058805-A2.  
15-JUL-2004.

25-DEC-2003; 2003WO-JP016715.  
XX

PR 26-DEC-2002; 2002JP-00376365.  
 PR 27-DEC-2002; 2002US-0436473P.  
 PR 25-APR-2003; 2003JP-00122113.  
 PR 28-APR-2003; 2003US-0465792P.  
 PR 21-OCT-2003; 2003JP-00360559.  
 PR 22-OCT-2003; 2003US-0512846P.  
 XX (ASAH-) ASahi KASEI PHARMA CORP.  
 PA Matsuda A, Yoneta S;  
 XX  
 XX WPI; 2004-593134/57.  
 DR P-PSDB; ADQ95918.  
 XX  
 PT New purified protein involved in T cell activation, useful for  
 PT diagnosing, preventing and/or treating acquired immunodeficiency  
 PT syndrome, autoimmune (e.g. rheumatoid arthritis, and diabetes), allergic  
 PT and infectious diseases.  
 XX  
 XX Claim 4; SEQ ID NO 95; 2828pp; English.  
 XX  
 CC The invention relates to purified proteins and genes encoding them, that  
 CC are involved in T cell activation (I) and has an amino acid deletion,  
 CC substitution or addition in the amino acid sequences. The methods and  
 CC compositions of the present invention are useful for the diagnosis,  
 CC prevention and/or treatment of autoimmune disease (rheumatoid arthritis,  
 CC asthma, multiple sclerosis and diabetes), allergic disease, infectious  
 CC disease, AIDS, and acute or chronic rejection at organ transplant or bone  
 CC marrow transplant. This sequence corresponds to a cDNA for a protein  
 CC involved in T cell activation.  
 XX  
 SQ Sequence 3302 BP; 980 A; 648 C; 773 G; 901 T; 0 U; 0 Other;  
 XX

Alignment Scores:  
 Pred. No.: 0 Length: 3302  
 Score: 754.00 Matches: 948  
 Percent Similarity: 98.54% Conservative: 0  
 Best Local Similarity: 98.54% Mismatches: 1  
 Query Match: 79.45% Indels: 14  
 DB: 12 Gaps: 0

US-09-671-687A-3 (1-949) x ADQ95917 (1-3302)  
 QY 1 MetSerSerGlyLeuThrSerGlnGluLysValThrSerProTyrTrpGluGluArgile 20  
 DB 243 ATGAGTTTCAGCTTATGGAGCCAGAAAAGTCACTTCACCTCTCTGGAGAGCGGATT 302  
 QY 21 PheTyrLeuLeuLeuGlnGluCysSerValThrAspLysGlnThrGlnLysLeuLeuLys 40  
 DB 303 TTTTACTTGTCTTCAAGATGTCAGCGTTACAGACAAACAAACAAAGCTCTTAAA 362  
 QY 41 ValProLysGlySerIleGlyGlnTyrIleGlnAspArgSerValGlyHisSerArgile 60  
 DB 363 GTACCGAAGGGAAGTATAGGACAGTATATTCAAGATCGTTCTGTGGGCGATTCAAGGATT 422  
 QY 61 ProSerAlaLysGlyLysLysAsnGlnIleGlyLeuLysIleLeuGluGlnProHisala 80  
 DB 423 CCTTCTGCAAAAGGCAAGAAATCAGATTGGATTGAAATTAATAATCTAGAGCAACCTCATGCA 482  
 QY 81 ValLeuPheValAspGlu---AspValValGluIleAsnGluLysPheThrGluLeuLeu 99  
 DB 483 GTTCTCTTTGTGATGAAGAGTGTGTAGAGATAAATGAAGAAGTTACAGAGTTACTT 542  
 QY 100 LeuAlaIleThrAsnGluGluArgPheSerLeuPheLysAsnArgAsnArgLeuSer 119  
 DB 543 TTGGCAATTACCAATTTGTAGGAGAGGTTACGCTGTTTAAAAACAGAAACAGACTAAGT 602  
 QY 120 LysGlyLeuGlnIleAspValGlyCysProValLysValGlnLeuArgSerGlyGluGlu 139  
 DB 603 AAAGGCTCCCAATAGACGTGGGCTGTCTGTGAAGATACAGCTCAGATCTGGGGAGAA 662  
 QY 140 LysPheProGlyValValArgPheArgGlyProLeuLeuAlaGluArgThrValSerGly 159

DB 663 AAATTTCTCGAGTTGTAGCTTTCAGAGGACCCCTGTTAGCAGAGGAGACAGTCTCCGA 722  
 QY 160 IlePhePheGlyValGluLeuLeuGluGlyArgGlyGlnGlyPheThrAspGlyVal 179  
 DB 723 ATATTTCTTTGGATTGAATTTGCTGGAAGAGTCTGTGTCAGAGTTTCTACTGACGGGGTG 782  
 QY 180 TyrGlnGlyLysGlnLeuPheGlnCysAspGluAspCysGly---PheValAlaLeuAsp 198  
 DB 783 TACCAAGGGAAACAGCTTTTTCAGTGTGATGAAGATTGTGGCGTGTGTTGTCATTGGAC 842  
 QY 199 LysLeuGluLeuIleGluAspAspThrAlaLeuGluSerAspTyrAlaGlyProGly 218  
 DB 843 AAGCTAGAACTCATAGAAGATGATGACACTGCAATTGGAAGATGATTACGACAGTCTCTGG 902  
 QY 219 AspThrMetGlnValGluLeuProLeuGluIleAsnSerArgValSerLeuLysGly 238  
 DB 903 GACAAATGCGAGTCCGAATCTCTCTTTGGAAATAAACTCCAGAGTTTCTTTGAA-GGT 961  
 QY 239 -GlyGluThrIleGluSerGlyThrValIlePheCysAspValLeuProGlyLysGluSe 258  
 DB 962 TGGAGAAACAATAGAAATCTGGAACAGTTATATTCTGTGATGTTTTCAGAGGAAAGAAAG 1021  
 QY 258 IleuGlyTyrPheValGlyValAspMetAspAsnProIleGlyAsnTrpAspGlyArgph 278  
 DB 1022 CTTAGGATATTTTGTGTGTGGACATGATAACCTATTGGCAACTGGGATGGAAGATT 1081  
 QY 278 eAspGlyVal---LeuCysSerPheAlaCysValGluSerThrIleLeuLeuHisIleSe 297  
 DB 1082 TGATGGAGTGCAGCTTTGTAGTFTTTCGCTGTGTTGAAAGTAGCAATTTCTATTGCACATCA 1141  
 QY 297 nAspIleIleProGluSerValThrGlnGluArgProProLysLeuAlaPheMetSe 317  
 DB 1142 TGATATCATCCAGAGAGTGTGACCGAGGAAGAGGCGCTCCCAACTTGCCTTTATGTC 1201  
 QY 317 rArgGlyValGlyAspLysGlySerSerHisAsnLysProLysAlaThrGlySerTh 337  
 DB 1202 AAGAGTGTGTGGGACAAAGGTTTCATCCAGTCATAATAAACCAAGGCTACAGGATCTAC 1261  
 QY 337 rSerAspProGlyAsnArg---ArgSerGluLeuPheTyrThrLeuAsnGlySerSerVa 356  
 DB 1262 CTCAGACCTCGAAATAGAAACAGATCTGAATTTATTTATACCTTAAATGGGTCTCTGT 1321  
 QY 356 lAspSerGlnProGlnSerLysSerLysAsnThrTrpTyrIleAspGluValAlaGluAs 376  
 DB 1322 TGACTCACACACCAATCAAAATCAAAATATACATGTTGATGAAGTTGCAAGAA 1381  
 QY 376 pProAlaLysSerLeuThrGluIleSerThrAspPheAspArgSerSerProLeuGlu 396  
 DB 1382 CCCTGCAAAATCTCTTTACAGAGATATCTACAGACTTTTGACCGTCTTCCACCACCTCCA 1441  
 QY 396 nProProValAsnSerLeuThrThrGluAsnArgPheHisSerLeuProPheSerIle 416  
 DB 1442 GCCTCTCTCTGTGAATCACTGACCACCGAGAACAGATTCCACTCTTTACCAATTCACTCT 1501  
 QY 416 uThrLysMetProAsnThrAsnGlySerIleGlyHisSerProLeuSerLeuSerAlaG 436  
 DB 1502 CACCAGATGCCCCAATACCAATGGAAGATTGCGCACAGTCCACTTTCTCTGTGAGCCCA 1561  
 QY 436 nSerValMetGluGluLeuAsnThrAlaProValGlnGluSerProProLeuAlaMetPr 456  
 DB 1562 GTCTGTATGGAAGAGCTTAACACTGACCCGCTCCAGAGAGTCCACCTTGGCCCATGCC 1621  
 QY 456 oProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLysGluAsnProPr 476  
 DB 1622 TCCTGGAACTCACATGCTTAGAAGTGGGCTCATTTGGCTGAAGTTAAGGAGAACCTCC 1681  
 QY 476 oPheTyrGlyValIleArgTropIleGlyGlnProProGlyLeuAsnGluValLeuAlaG 496  
 DB 1682 TTTCTATGGGGTAATCCGTTGGATCGGTGAGCCAGGACTGAATGAAGTCTCGCTGG 1741  
 QY 496 yLeuGluLeuGluAspGluCysAlaGlyCysThrAspGlyThrPheArgGlyThrArgTy 516  
 DB 1742 ACTGGAATCTGGAAGATGAGTGTGAGGCTGTACGGATGGAACCTTTCAGAGGCACTCGGTA 1801



QY 516 rPheThrCysalaLeuLysLysAlaLeuPheValLysLeuLysSerCysArgProAspSe 536  
DB 1802 TTTCACTGTGCGCTGAAGAAGCGCTGTTGTGAAACTGAAGAGCTGCAGGCGCTGACTC 1861  
QY 536 rArgPheAlaSerLeuGlnProValSerAsnGlnIleGluArgCysAsnSerLeuAlaPh 556  
DB 1862 TAGGTTTCATCATTCGACCGCGTTTCCAATCAGATTGAGCGCTGTAACCTCTTTAGCAT 1921  
QY 556 eGlyGlyTyrrLeuSerGluValValGluGluLeuThrProProLysMetGluLysGluG 576  
DB 1922 TGGAGGCTACTTAAAGTGAAGTAGTAGAAGAAATACTCCACCAAAATGGAAGAAAGG 1981  
QY 576 yLeuGluIleMetIleGlyLysLysGlyIleGlnGlyHisTyrrAsnSerCysTyrrLe 596  
DB 1982 CTTGGAGATAATGATTGGGAAGAAAGGCAATCCAGGTCATTACAAATCTTTGTACTT 2041  
QY 596 uAsnSerThrLeuPheCysLeuPheAlaPheSerValLeuAspThrValLeuLeuAr 616  
DB 2042 AGACTCAACCTTATTCGTCTTATTTGCTTTAGTTCTGTTCTGGACACTGTGTTACTTAG 2101  
QY 616 gProLysGluLysAsnAspValGluTyrrTyrrSerGluThrGlnGluLeuArgThrGl 636  
DB 2102 ACCCAAGAAAGAAAGCAATGATAGATATATATATGTAAGCCCAAGAGCTACTGAGGACAG 2161  
QY 636 uIleValAsnProLeuArgIleTyrrGlyTyrrValCysAlaThrLysIleMetLysLeuAr 656  
DB 2162 AATTGTTAATCTCTGAGAATATATGGAATGTGTGTCACCAAAATATGAAACTGAG 2221  
QY 656 gLysIleLeuGluLysValGluAlaSerGlyPheThrSerGluGluLysAspProGl 676  
DB 2222 GAAAATACTTGAAAAGGTGGAGCTGCATCAGGATTTACCTCTGAAGAAAAGATCCTGA 2281  
QY 676 uGluPheLeuAsnIleLeuPheHisIleLeuArgValGluProLeuLysIleAr 696  
DB 2282 GGAATTTCTTGATATTTCTGTTTCATCATATATTTTAAGGGTAGAACCTTTTGTAAAAATAAG 2341  
QY 696 gSerAlaGlyGlnLysValGlnAspCysTyrrPheTyrrGlnIlePheMetGluLysAsnGl 716  
DB 2342 ATCAGCAGGTCAAAAGGTACAGATTTGTTACTTCTATCAATTTTATGGAAAAAATGA 2401  
QY 716 uLysValGlyValProThrIleGlnGlnLeuLeuGluTrpSerPheIleAsnSerAsnLe 736  
DB 2402 GAAGTTGGCGTTCACCAATTCACAGTTGTTAGATGGTCTTTTATCAACAGTAACCT 2461  
QY 736 uLysPheAlaGluAlaProSerCysLeuIleIleGlnMetProArgPheGlyLysAspPh 756  
DB 2462 GAAATTTGCAGAGGCACCATCATGCTGATTATTCAGATGCCTCGATTTTGGAAAAGACTT 2521  
QY 756 eLysLeuPheLysLysIlePheProSerLeuGluLeuAsnIleThrAspLeuLeuGluAs 776  
DB 2522 TAAACTATTTAAAAAAATTTTTCCTCTCTGGAATTAATAATAACAGATTTCATCTTGAAGA 2581  
QY 776 pThrProArgGlnCysArgIleCysGlyGlyLeuAlaMetTyrrGluCysArgGluCysTy 796  
DB 2582 CACTCCACAGACGTCGCCGATATGAGAGGCGTTGCAATGTATGAGTGTAGAGAAATGCTA 2641  
QY 796 rAspAspProAspIleSerAlaGlyLysIleLysGlnPheCysLysThrCysAsnThrGl 816  
DB 2642 CGACGATCCGACATCTCAGCTGGAAAATCAAGCAGTTTGTAAAACTGCAACACTCA 2701  
QY 816 nValHisLeuHisProLysArgLeuAsnHisLysTyrrAsnProValSerLeuProLysAs 836  
DB 2702 AGTCCACCTTCATCCGAAGAGCTGAATCATATAATATATACCCAGATGTCTATTTCCCAAGA 2761  
QY 836 pLeuProAspTrpAspTrpArgHisGlyCysIleProCysGlnAsnMetGluLeuPheAl 856  
DB 2762 CTTACCCGACTGGGACTGGAGACCGCTGCATCCCTTCCCAAGAAATATGAGTATTTGTC 2821  
QY 856 aValLeuCysIleGluThrSerHisTyrrValAlaPheValLysTyrrGlyLysAspAspSe 876  
DB 2822 TGTCTCTGTCATAGAAACAGCCACTATGTTCTCTTGTGAGTATGGAAGACCATTC 2881

QY 876 rLaTrpLeuPhePheAspSerMetAlaAspArgAspGlyGlyGlnAsnGlyPheAsnIl 896  
DB 2882 TGCTGGCTCTTCTTTGACAGCATGGCCGATCGGATGGTGGTCAGAAATGGCTTCAACAT 2941  
QY 896 eProGlnValThrProCysProGluValGlyGluTyrrLeuLysMetSerLeuGluAspLe 916  
DB 2942 TCTCAAGTCACCCCATGCCAAGTAGGAGAGTACTTGAAGATGTCTCTGGAAGACCT 3001  
QY 916 uHisSerLeuAspSerArgArgIleGlnGlyCysAlaArgArgLeuLeuCysAspAlaTy 936  
DB 3002 GCATTCCTTGGACTCCAGAGAAATCCAGGCTGTGCAGAGACTGCTTGTGTATGCATA 3061  
QY 936 rMetCysMetTyrrGlnSerProThrMetSerLeuTyrrLys 949  
DB 3062 TATGTCATGTACCAGATCCCAACATGATGTTGTACAA 3101  
RESULT 2  
ADC24898  
ID ADC24898 standard; cDNA; 6831 BP.  
XX  
AC ADC24898;  
XX  
DT 18-DEC-2003 (first entry)  
XX  
DE Human breast specific nucleic acid (BSNA) DEX0238\_49, SEQ ID NO:49.  
XX  
KW Human; breast specific polypeptide; BSP; breast specific nucleic acid;  
KW BSNA; breast cancer; metastasis; non-cancerous disease; breast tissue;  
KW identification; monitoring; diagnosis;  
KW engineered breast tissue production; transgenic animal; drug screening;  
KW cytostatic; gene therapy; vaccine; chromosome 16p13.3; gene; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO2003020900-A2.  
XX  
PD 13-MAR-2003.  
XX  
PF 29-AUG-2002; 2002WO-US027777.  
XX  
PR 31-AUG-2001; 2001US-0316306P.  
XX  
PA (DIAD-) DIADEXUS INC.  
XX  
PI Sun Y, Liu C, Salceda S;  
XX  
DR WPI; 2003-290182/28.  
DR P-PSDB; ADC24816.  
XX  
PT New breast specific polypeptide useful for identifying, diagnosing,  
PT monitoring, staging, imaging and treating breast cancer and non-cancerous  
PT disease states in breast.  
XX  
PS Example 1; SEQ ID NO 49; 264pp; English.  
XX  
CC The invention relates to breast specific polypeptides (BSNs) and nucleic  
CC acids (breast specific nucleic acids; BSNAs) encoding them. The invention  
CC also relates to vectors and host cells comprising a BSN sequence;  
CC antibodies against BSNs; the recombinant production of BSNs; methods of  
CC detection of BSNAs or BSNs in a sample; kits for detecting a risk of  
CC cancer or presence of cancer in a patient; and vaccines comprising a BSN  
CC or BSP. The invention additionally discloses fragments, mutants, fusion  
CC proteins, homologous proteins and allelic variants of BSNs; methods for  
CC identifying and designing agonists and antagonists of BSNs; methods for  
CC identifying and monitoring breast tissue; producing engineered breast  
CC tissue for treatment and research; producing transgenic animals and cells  
CC comprising BSN sequences; aptamers evolved to bind specifically to BSNs;  
CC and single exon probes based on BSN sequences. BSNs, BSNAs and  
CC antibodies against BSNs are useful for identifying, diagnosing,  
CC monitoring, staging, imaging and treating breast cancer (including breast  
CC cancer metastases), and non-cancerous disease states in breast tissue.  
CC BSNs and BSNAs may additionally be used to identify and monitor breast  
CC tissue, in screening for BSP agonists and antagonists, and in the

CC production of engineered breast tissue for treatment or research. BSNAs  
 CC may also be used in gene therapy and in the production of transgenic  
 CC animals and cells. The present sequence represents a breast specific  
 CC nucleic acid (BSNA) used in an example of the invention.

XX Sequence 6831 BP; 1987 A; 1256 C; 1494 G; 2094 T; 0 U; 0 Other;

#### Alignment Scores:

Pred. No.: 0 Length: 6831  
 Score: 754.00 Matches: 948  
 Percent Similarity: 98.54% Conservative: 0  
 Best Local Similarity: 98.54% Mismatches: 1  
 Query Match: 79.45% Indels: 14  
 DB: 10 Gaps: 0

US-09-671-687A-3 (1-949) x ADC24898 (1-6831)

QY 1 MetSerSerGlyLeuTrpSerGlnGluLysValThrSerProTyTrpGluuArgile 20  
 Db 1585 ATGAGTTTCAGGCTTATGGAGCCCAAGAAAAAGTCACTTCACCCCTACTGGGAAGAGCGGATT 1644  
 QY 21 PheTyrlLeuLeuGlnGluCysSerValThrAspLysGlnThrGlnLysLeuLeuLys 40  
 Db 1645 TTTTACTTGCCTTCAGAAATGCGAGCGTTACAGACAAACAAACAAAGCTCTCTAAA 1704  
 QY 41 ValProLysGlySerIleGlyGlnTrpIleGlnAspArgSerValGlyHisSerArgile 60  
 Db 1705 GTACCGAGGGAAGTATAGGACAGTATATCAAGATCGTTCTGTGGGCAATTCAGGATT 1764  
 QY 61 ProSerAlaLysGlyLysGlnGlnIleGlyLeuLysIleLeuGluGlnProHisAla 80  
 Db 1765 CCTTCTGCAAAAGGCAAGAAAAATCAGATTGGATTAAAAATTTCTAGAGCAACCTCATGCA 1824  
 QY 81 ValLeuPheValAspGlu---AspValValGluIleAsnGluLysPheThrGluLeuLeu 99  
 Db 1825 GTTCTCTTGTGTGAAAGAGATGTTGTAGAGATAAATGAAAGTTTCACAGATTACTT 1884  
 QY 100 LeuAlaIleThrAsnGlyGluGluArgPheSerLeuPheLysAsnArgAsnArgLeuSer 119  
 Db 1885 TTGGCAATTACCAATTGTGAGGAGAGTTTCAGCTGTTTAAAAACAGAAACAGACTNAGT 1944  
 QY 120 LysGlyLeuGlnIleAspValGlyCysProValLysValGlnLeuArgSerGlyGluGlu 139  
 Db 1945 AAAGGCTTCCAAATAGACGTGGGCTGCTCTGTGAAAGTACAGCTCAGATCTGGGGAAGA 2004  
 QY 140 LysPheProGlyValValArgPheArgGlyProLeuLeuAlaGluArgThrValSerGly 159  
 Db 2005 AAATTTCTGGAGTTGTACGCTTCAGAGGACCCCTGTTAGCAGAGAGACAGCTCTCCGA 2064  
 QY 160 IlePhePheGlyValGluLeuLeuGluGluArgGlyArgGlyGlnGlyPheThrAspGlyVal 179  
 Db 2065 ATATCTTTGGAGTTGAATTCGTAAGAGAGTCTGTGTCAAGGTTTTCACCTACCGGGTG 2124  
 QY 180 TyrGlnGlyLysGlnLeuPheGlnCysAspGluAspCysGly---PheValAlaLeuAsp 198  
 Db 2125 TACCAAGGGAACAGCTTTTTCAGTGTGATGAAGATTGTGGCGTGTGTTGTCATTGGAC 2184  
 QY 199 LysLeuGluLeuIleGluAspAspThrAlaLeuGluSerAspTyAlaGlyProGly 218  
 Db 2185 AAGCTAGAACTCATAGAGATGATGACACTGCAATTGGAAGTGAATTACGACAGTCTCTGG 2244  
 QY 219 AspThrMetGlnValGluLeuProProLeuGluIleAsnSerArgValSerLeuLysGly 238  
 Db 2245 GACCAATGTCAGGTCGAATCTCTCTTTGGAAATAAACTCCAGAGTTTCTTTGAA -GGT 2303  
 QY 239 -GlyGluThrIleGluSerGlyThrValIlePheCysAspValLeuProGlyLysGluSe 258  
 Db 2304 TGGGAACAATAGATCTGGACAGTTATATCTGTGATGTTTTCACGAGGAAAGAAAG 2363  
 QY 258 rLeuGlyTyrPheValGlyValAspMetAspAsnProIleGlyAsnTrpAspGlyArgph 278  
 Db 2364 CTAGGATATTTTGTGGTGGACATGATGAACCTATTGGCAACTTGGGATGGAAGATT 2423

QY 278 eAspGlyVal---LeuCysSerPheAlaCysValGluSerThrIleLeuLeuHisIleLe 297  
 Db 2424 TGATCGAGTGCAGCTTTTGTAGTTTTCGCTGTGTTGAAAGTACAATCTTATTGCACATCAA 2483  
 QY 297 nAspIleIleProGluSerValThrGlnGluArgArgProProLysLeuAlaPheMetSe 317  
 Db 2484 TGATATCATCCAGAGAGTGTGACGAGGAAGAGGAGGCTCCCAAACTTGCCCTTTATGTC 2543  
 QY 317 rArgGlyValGlyAspLysGlySerSerHisAsnLysProLysAlaThrGlySerTh 337  
 Db 2544 HAGAGTGTGGGGACAAAGGTTTCATCCAGTATATAAACAAGGCTCAGAGATCTAC 2603  
 QY 337 rSerAspProGlyAsnArg---ArgSerGluLeuPheTyThrLeuAsnGlySerSerVa 356  
 Db 2604 CTCAGACCTCGAAATAGAAACAGATCTGAATTTATATATCTTAAATGGGTCTTCTGT 2663  
 QY 356 lAspSerGlnProGlnSerLysSerLysAsnThrTrpTyrlleAspGluValAlaGluAs 376  
 Db 2664 TGACTCACACCAACCAATCAAAATCAAAATAATACATGGTACATGATGAAGTTGCAAGA 2723  
 QY 376 pProAlaLysSerLeuThrGluIleSerThrAspPheAspArgSerSerProLeuGlu 396  
 Db 2724 CCTCTCAAAATCTCTTACAGAGATATCTACAGACTTTGACCGTTCTTACACCACTCCA 2783  
 QY 396 nProProValAsnSerLeuThrThrGluAsnArgPheHisSerLeuProPheSerLe 416  
 Db 2784 GCCTCTCTGTGAATCTCACTGACCAACGAGAACAGATTCCACTCTTTACCATTCAGTCT 2843  
 QY 416 uThrLysMetProAsnThrAsnGlySerIleGlyHisSerProLeuSerLeuSerAlaGl 436  
 Db 2844 CACCAAGATGCCCAATACCAATGGAAGTATTGGCCACACAGTCCACTTTCTCTGTGAGCCA 2903  
 QY 436 nSerValMetGluGluLeuAsnThrAlaProValGlnGluSerProProLeuAlaMetPr 456  
 Db 2904 GTCTTAATGGAAGAGCTAAACACATGACCCGCTCCAGAGAGTCCACCCCTTGCCCATGCC 2963  
 QY 456 pProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLysGluAsnProPr 476  
 Db 2964 TCCTGGGAACCTCACATGGTCTAGAAAGTGGGCTCATTTGGCTGAAGTTTAGGAGAACCTCC 3023  
 QY 476 oPheTyrlGlyValIleArgTrpIleGlyGlnProProGlyLeuAsnGluValLeuAlaGl 496  
 Db 3024 TTTCTATGGGGTAATCCGTTGGATCGGTGAGCTGAGTGAATGAAGTCTGCTGCTGG 3083  
 QY 496 yLeuGluLeuGluAspGluCysAlaGlyCysThrAspGlyThrPheArgGlyThrArgTy 516  
 Db 3084 ACTGGAACCTGGAAGATGAGTGTGCAAGGCTGTGAGGATGGAACCTTTACAGAGGCACTCGGTA 3143  
 QY 516 rPheThrCysAlaLeuLysLysAlaLeuPheValLysLeuLysSerCysArgProAspSe 536  
 Db 3144 TTTCACTGTGCCCTGGAAGAGCGCTGTTTGTGAACCTGGAAGAGCTGAGGCGCTGACTC 3203  
 QY 536 rArgPheAlaSerLeuGlnProValSerAsnGlnIleGluArgCysAsnSerLeuAlaPh 556  
 Db 3204 TAGTGTTCATCATTCGACCGCGTTTCCAATCAGATTGAGCGCTGTAACTCTTTAGCATTT 3263  
 QY 556 eGlyGlyTyrLeuSerGluValValGluGluAsnThrProProLysMetGluLysGluGl 576  
 Db 3264 TGGAGGCTACTTAACTGAAAGTAGTAGAAGAAAATACTCCACCAAAATGGAAGAAAGG 3323  
 QY 576 yLeuGluIleMetIleGlyLysLysGlyIleGlnGlyHisTyAsnSerCysTyrlle 596  
 Db 3324 CTTGAGAGATAATGATTGGGAAGAAAGGCAATCCAGGCTCATTAACAATTTCTTGTACTTT 3383  
 QY 596 uAspSerThrLeuPheCysLeuPheAlaPheSerSerValLeuAspThrValLeuLeuAr 616  
 Db 3384 AGACTCAACCTTATTCGCTTATTTGCTTATTTGCTTGTCTGGACACTGTGTACTTAG 3443  
 QY 616 pProLysGluLysAsnAspValGluTyTrpSerGluThrGlnGluLeuLeuArgThrGl 636  
 Db 3444 ACCCAAGAAAGAACGATGAGATAATATATAGTGAACCCCAAGAGCTACTGAGGACAGA 3503  
 QY 636 uIleValAsnProLeuArgIleTyGlyTyTrValCysAlaThrLysIleMetLysLeuAr 656

Db 3504 AATTGTTAATCTCTGAGAAATATATGATATGTTGTCACAAAAAATATGAAACTGAG 3563  
Qy 656 GLysIleLeuGluLysValGluAlaAaSerGlyPheThrSerGluGluLysAspProG1 676  
Db 3564 GAAATACCTTGAAGAAGTGGAGCTGCATCAGATTCTCTGGAAGAAAGATCCTGA 3623  
Qy 676 uGluPheLeuAsnIleLeuPheHisIleLeuArgValGluProLeuLeuLysIleAr 696  
Db 3624 GGAATTCITGAATATCTGTTTCATCATATTTTAAGGGTAGAACCTTTTGTAAATAAG 3683  
Qy 696 gSerAlaGlyGlnLysValGlnAspCysTyrPheTyrGlnIlePheMetGluLysAsnG1 716  
Db 3684 ATCAGCAGCTCAAAAGGTACAAGATTGTTACTTCTATCAAAATTTTATGAAAAAATGA 3743  
Qy 716 uLysValGlyValProThrIleGlnLeuLeuGluTyrSerPheIleAsnSerAsnLe 736  
Db 3744 GAAAGTTGGCGTTCACCAATTCAGCAGTGTGTAGAAATGGTCTTTTATCAACAGTAACCT 3803  
Qy 736 uLysPheAlaGluAlaProSerCysLeuIleIleGlnMetProArgPheGlyLysAspPh 756  
Db 3804 GAAATTTGCAGAGGCACCATCATGCTGATTTACAGATGCCCTCGATTTGGAAAAAGACTT 3863  
Qy 756 eLysLeuPheLysIlePheProSerLeuGluLeuAsnIleThrAspLeuLeuGluAs 776  
Db 3864 TAAACTATTTAAAAAATTTTCTCTCTGGAATTAATATAACAGATTACTTGAAGA 3923  
Qy 776 pThrProArgGlnCysArgIleCysGlyLeuAlaMetTyrGluCysArgGluCysTy 796  
Db 3924 CACTCCAGACAGTCCGAGATATGGAGGGTTCGAATGTATGAGTGTAGAGAAATGCTA 3983  
Qy 796 rAspAspProAspIleSerAlaGlyLysIleLysGlnPheCysLysThrCysAsnThrG1 816  
Db 3984 CGACATCCGGACATCTACGCTGGAAAAATCAAGCAGTTTGTAAAACTTCACACTCA 4043  
Qy 816 nValHisLeuHisProLysArgLeuAsnHisLysTyrAsnProValSerLeuProLysAs 836  
Db 4044 AGTCCACCTTCATCCGAGAGCTGAATCATAATATACCCAGTGTCTCTCCCAAGA 4103  
Qy 836 pLeuProAspTrpAspTrpArgHisGlyCysIleProCysGlnAsnMetGluLeuPheAl 856  
Db 4104 CTTACCCGACTGGGACTGGAGACACGGCTGCATCCCTGCCAGATATGGAGTTATTTC 4163  
Qy 856 aValLeuCysIleGluThrSerHisTyrValAlaPheValLysTyrGlyLysAspSpSe 876  
Db 4164 TGTTCTCTGCATAGAAAACCAAGCCACTATGTCTTTGTGAAGTATGGGAAGACGATTTC 4223  
Qy 876 rAlaTrpLeuPhePheAspSerMetAlaAspArgAspGlyGlyGlnAsnGlyPheAsnI1 896  
Db 4224 TGCCTGGCTCTCTTTTGACAGATGCGCGATCGGGATGGTGTGAGAAATGGCTTCAACAT 4283  
Qy 896 eProGlnValThrProCysProGluValGlyGlyLeuLysMetSerLeuGluAspLe 916  
Db 4284 TCCTCAAGTCACCCCATGCCAGAGTAGGAGATGACTTGAAGATGCTCTGGAAGACCT 4343  
Qy 916 uHisSerLeuAspSerArgArgIleGlnGlyCysAlaArgArgLeuLeuCysAspAlaTy 936  
Db 4344 GCATTCTCTGGACTCCAGAGAAATCCAGGCTGTGCACGAAGACTCTTTGTGTATGCATA 4403  
Qy 936 rMetCysMetTyrGlnSerProThrMetSerLeuTyrLys 949  
Db 4404 TATGTGCATGTACCAAGATCCCAACAAATGATGTTTGTACAAA 4443

## RESULT 3

AAA08589  
ID AAA08589 standard; DNA; 4527 BP.  
XX  
AC AAA08589;  
XX  
DT 19-JUL-2000 (first entry)  
XX  
DE Human cytoskeleton associated protein 9 (CYSKP-9) coding sequence.  
XX

KW Cytoskeleton associated protein; CYSKP-9; cancer; proliferative;  
KW autoimmune; inflammatory, vesicle trafficking; neurological;  
KW cardiovascular; cell motility; reproductive; muscle disorder; as.

OS Homo sapiens.

XX Key Location/Qualifiers  
FH 926..3121  
FT CDS /tag= a  
FT /product= "CYSKP-9"

XX WO200017355-A2.

XX 30-MAR-2000.

XX 17-SEP-1999; 99WO-US021565.

XX 18-SEP-1998; 98US-0172226P.

XX 27-APR-1999; 99US-0131321P.

XX (INCY-) INCYTE PHARM INC.

PI Lal P, Tang YT, Yue H, Hillman JL, Bandman O, Corley NC;

PI Guegler KJ, Patterson C, Azimzai Y, Baughn MR;

DR WPI; 2000-283582/24.

DR P-PSDB; AA91954.

XX Human cytoskeleton associated proteins, used to treat cell proliferative,  
PT autoimmune/inflammatory, vesicle trafficking, neurological, cell  
PT motility, reproductive and muscle disorders.

PS Claim 9; Page 107-109; 113pp; English.

XX AAA08581-96 encode human cytoskeleton associated proteins 1 to 16 (CYSKP-  
CC 1 to CYSKP-16) respectively. The sequences can be used to treat and  
CC diagnose cancer and cell proliferative, autoimmune/inflammatory, vesicle  
CC trafficking, neurological, cardiovascular, cell motility, reproductive  
CC and muscle disorders. Pharmaceutical compositions containing CYSKP-2 to  
CC CYSKP-16 can be used to treat or prevent disorders associated with  
CC decreased expression or activity of CYSKP (claimed), for example,  
CC atherosclerosis, cirrhosis, hepatitis, myelofibrosis, psoriasis, cancers,  
CC autoimmune/antiinflammatory disorders such as allergies, anemia, asthma,  
CC acquired immunodeficiency syndrome (AIDS), Crohn's disease, diabetes  
CC mellitus, Goodpasture's syndrome, osteoarthritis, scleroderma and trauma.  
CC CYSKP antagonists can be used to treat or prevent a disorder associated  
CC with increased expression or activity of CYSKP (claimed)

XX SQ Sequence 4527 BP; 1351 A; 823 C; 1005 G; 1348 T; 0 U; 0 Other;

## Alignment Scores:

Pred. No.: 0 Length: 4527  
Score: 711.00 Matches: 944  
Percent Similarity: 98.44% Conservative: 0  
Best Local Similarity: 98.44% Mismatches: 2  
Query Match: 74.92% Indels: 15  
DB: 3 Gaps: 0

US-09-671-687A-3 (1-949) x AAA08589 (1-4527)

Qy 4 GlyLeuTrpSerGlnGluLysValThrSerProTyrTrpGluGluArgIlePheTyrLeu 23

Db 270 GGCTTATGTGAGCCAAAGAAAGTCCTTACCCTACTCTGGGAAGCGGATTTTACTTG 329

Qy 24 LeuLeuGlnGluCysSerValThrAspLysGlnThrGlnLysLeuLysValProLys 43

Db 330 CTTCTTCAGATGTCAGCGTTACAGACAAACACAAAGCTCTTAAGTACCGAA- 388

Qy 44 GlySerIleGlyGlnTyrIleGlnAspArgSerValGlyHisSerArgIleProSerAla 63

Db 389 GGAAGTATAGGACAGTATATTCAAGATCGTCTCTGTGGGCAATCAAGGATTCCTTGTGCA 448

Qy 64 LysGlyLysLysAsnGlnIleGlyLeuLysIleLeuGluGlnProHisAlaValLeuPhe 83

Db	449	AAAGCAAGAAAAATCAGATTGGATTAAAAATTCCTAGACCAACCTCATGCGAGTTCCTCTT	1528	GGCCAAATACCAATGGAAAGTATTGGCCACAGTCCACTTTCTCTGTGTCAGCCCACTGTGTAAT	1587
Qy	84	ValAspGlu---AspValValGluIleAsnGluLysPheThrGluLeuLeuLeuAlaIle	439	tGluGluLeuAsnThrAlaProValGlnGlnSerProProLeuAlaMetProProGlyAs	459
Db	509	GTTCGATGAAAGAGATGTTGAGAGATAAATGAAAGATTCACAGAGTACTTTGGCAATT	1588	GGAGAGCTAAACACTGCACCGTCCAAGAGATCCACCTTGGCCATGCCTCCTGGGAA	1647
Qy	103	ThrAsnCysGluGluArgPheSerLeuPheLysAsnArgAsnArgLeuSerLysGlyLeu	459	nSerHisGlyLeuGluValGlySerLeuAlaGluValLysGluAsnProProPheTyrGln	479
Db	569	ACCAATTGTGAGGAGAGGTTGAGCGTGTAAAAACAGAAACAGACTAAGTAAAGGCCTC	1648	CTCACATGGTCTAGAAAGTGGGCTCATTTGGCTGAAGTTAAGGAGAACCTCTCTCTTATGG	1707
Qy	123	GlnIleAspValGlyCysProValLysValGlnLeuArgSerGlyGluGluLysPhePro	479	yValIleArgTrrIleGlyGlnProProGlyLeuAsnGlnValLeuAlaGlyLeuGluLe	499
Db	629	CAATAGACGTGGGCTGTCCTGTGAAAGTACAGCTGAGATCTGGGGAAGAAAAATTCCT	1708	GGTAATTCGTTGGATCGGTTCAGCCACAGGACTGAATGAAGTGTCTGGCTGGAAT	1767
Qy	143	GlyValValArgPheArgGlyProLeuLeuAlaGluArgThrValSerGlyIlePhePhe	499	uGluAspGluCysAlaGlyCysThrAspGlyThrPheArgGlyThrArgTyrPheThrCys	519
Db	689	GGAGTTGTACCTTCAGAGGACCCCTGTAGCAGAGAGACAGTCTCCGGAATATCTTT	1768	GGAGATGAGTGTGCAGGCTGTACGGATGGAACTTTCAGAGGCACCTCGGTATTTACCTTG	1827
Qy	163	GlyValGluLeuLeuGluGluArgGlyGlnGlyPheThrAspGlyValTyrGlnGly	519	sAlaLeuLysLysAlaLeuPheValLysLeuLysSerCysArgProAspSerArgPheAl	539
Db	749	GGAGTTGAATTCGTGGAAGAGTTCGTGTCGAAGTTTCACTGACGGGTGTACCAAGGG	1828	TGCCCTGAAGAAGGCGCTGTTGTCAAACCTGAAGAGCTGCAGGCTGACTCTAGGTTTGC	1887
Qy	183	LysGlnLeuPheGlnCysAspGluAspCysGly---PheValAlaIleuAspLysLeuGlu	539	aSerLeuGlnProValSerAsnGlnIleGluArgCysAsnSerLeuAlaPheGlyGlyTyr	559
Db	809	AAACAGCTTTTTCAGTGTGATGAAGATTGTGGCGTGTGTTGTTCATTTGGACAAGCTAGAA	1888	ATCATTTGAGCGCGGTTTCCAAATCAGATTGAGCGCTGAACCTTTTAGCATTTTGGAGGCTA	1947
Qy	202	LeuIleGluAspAspThrAlaLeuGluSerAspTyrAlaGlyProGlyAspThrMet	559	rLeuSerGluValValGluGluAsnThrProProLysMetGluLysGluGlyLeuGluIrl	579
Db	869	CTCATAGAAGATGATGACACTGCATTGGAAGTGATTACGCAAGTCTCTGGGACCAATG	1948	CTTAAGTGAAGTAGTAGAAGAAATATCTCCACAAATATGGAAAAGAGGCTTGGAGAT	2007
Qy	222	GlnValGluLeuProProLeuGluIleAsnSerArgValSerLeuLysGly-GlyGluTh	579	eMetIleGlyLysLysLysGlyIleGlnGlyHisTyrAsnSerCysTyrLeuAspSerTh	599
Db	929	CAGGTGGAACCTCTCTTGGAAATAAATCCACAGATTCTTTGAA-GGTTGGAGAAAC	2008	AATGATTGGGAAGAAAGGACATCCAGGTCATTACAATCTCTGTACTTAGACTCAAC	2067
Qy	241	rIleGluSerGlyThrValIlePheCysAspValLeuProGlyLysGluSerLeuGlyTyr	599	rLeuPheCysLeuPheAlaPheSerSerValLeuAspThrValLeuLeuArgProLysGln	619
Db	988	AATAGAATCTGGAAACAGTATATTTCTGTCATGTTTGGCAGGAAAGAAAGCTTAGGATA	2068	CTTATTCCTGCTTATTTGCTTTTGTGTTCTGTTCTGACACTGTGTGTACTTAGACCCAAAGA	2127
Qy	261	rPheValGlyValAspMetAspAsnProIleGlyAsnTrpAspGlyArgPheAspGlyVa	619	uLysAsnAspValGluTyrTyrSerGluThrGlnGluLeuLeuArgThrGluIleValAs	639
Db	1048	TTTTGTTGGTGGACATGGATACCCCTATTGGCACTGGGATGGGAAGATTTGATGGAGT	2128	AAAGAACGATGTAGAATATATTAGTGAACCCCAAGACTACTGAGGACAGAAATTTGTTAA	2187
Qy	281	l---LeuCysSerPheAlaCysValGluSerThrIleLeuLeuHisIleAsnAspIleI	639	nProLeuArgIleTyrGlyTyrValCysAlaThrLysIleMetLysLeuArgLysIleLe	659
Db	1108	GCAGCTTTGTAGTTTGGTGTGTTGAAAGTACAATCTATTGCACATCAATGATATCAT	2188	TCCTCTGAGATATATGATATGTTGTGTCACCAAAATATTATGAACACTGAGGAAATACT	2247
Qy	300	eProGluSerValThrGlnGluArgArgProProLysLeuAlaPheMetSerArgGlyVa	659	uGluLysValGluAlaAlaSerGlyPheThrSerGluGluLysAspProGluGluPheLe	679
Db	1168	CCACAGAGTGTGACGCAAGAAAGGAGGCTCCCAAACTTGCCTTTATGTCAAGAGGTGT	2248	TGAAAAGGTGGAGGCTGCATCAGGATTTACCTCTGAAGAAAAGATCTCTGAGGAATCTT	2307
Qy	320	lGlyAspLysGlySerSerSerHisAsnLysProLysAlaThrGlySerThrSerAspPr	679	uAsnIleLeuPheHisIleLeuArgValGluProLeuLeuLysIleArgSerAlaGln	699
Db	1228	TGGGACAAAGGTTTCATCCAGTCATTAATAACCAAGGCTACAGGATCTACCTCAGACCC	2308	GAATATTCCTGTTTCATCATATTTTAAGGTTAGAACCCTTTGTCTAAAATAAAGTACAGCAG	2367
Qy	340	oGlyAsnArg---ArgSerGluLeuPheTyrThrLeuAsnGlySerSerValAspSerGln	699	yGlnLysValGlnAspCysTyrPheTyrGlnIlePheMetGluLysAsnGluLysValGln	719
Db	1288	TGGAAATAGAACACAGATCTGAATTTATTTATACCTTAAATGGGTCTTCTGTGACTACA	2368	TCAAAAGGTACAAAGATTGTACTTCTATCAAAATTTTATGGAATAAATGAGAAAGTTGG	2427
Qy	359	nProGlnSerLysSerLysAsnThrTrpTyrIleAspGluValAlaGluAspProAlaLys	719	yValProThrIleGlnGlnLeuLeuGluTrpSerPheIleAsnSerAsnLeuLysPheAl	739
Db	1348	ACCACAAATCCAAATCAAAAATACATGTTACATTCATGATGAGTTGAGAGACCTTGCAAA	2428	CGTTCACCAATTCAGCAGTGTGTAGATGGTCTTTTATCAACAGTAACTGAAATTTGC	2487
Qy	379	sSerLeuThrGluIleSerThrAspPheAspArgSerProProLeuGlnProProPr	739	aGluAlaProSerCysLeuIleIleGlnMetProArgPheGlyLysAspPheLysLeuPh	759
Db	1408	ATCTCTTACAGAGATATCTACAGACTTTGACCGTCTTCCACCACTCCAGCTCCTCC	2488	AGAGCAGCACCATCATGCTGATTTATTCAGATGCTCGATTTGGAAAAGACTTTTAAACTAT	2547
Qy	399	oValAsnSerLeuThrThrGluAsnArgPheHisSerLeuProPheSerLeuThrIysMe	759	eLysLysIlePheProSerLeuLeuAsnIleThrIlePheLeuLeuGluAspThrProAr	779
Db	1468	TGTGAACCTACGTACCCGAGAAACAGATTTCACCTTTTACCATTCAGCTCACCAGAT	2548	TAAAAAATTTTTTCTCTCTGGAATTAATAATAACAGATTTTACTTTGAAGACACTCCAG	2607
Qy	419	tProAsnThrAsnGlySerIleGlyHisSerProLeuSerLeuSerAlaGlnSerValMe	779	gGlnCysArgIleCysGlyGlyLeuAlaMetTyrGluCysArgGluCysTyrAspAspPr	799
			2608	ACAGTGGCGGATATGTGGAGGCTTGCATGTATGAGTGTAGAGAAATGCTACGACGATCC	2667

```
QY 799 oAspIleSerAlaGlyLysIleGlnPheCysLysThrCysAsnThrGlnValHisLe 819
Db 2668 GGACATCTCAGCTGGAAAAAATCAAGCAGTTTGTAAAAAAGCTGCAACACTCAAGTCCACCT 2727
QY 819 uHisProLysArgLeuAsnHisLysTyrAsnProValSerLeuProLysAspLeuProAs 839
Db 2728 TCATCCGAAGAGGCTGAATCATATAATATAACCCAGTGTCACTTCCCAAGACTTACCCGA 2787
QY 839 pTyrAspTrpArgHisGlyCysIleProCysGlnAsnMetGluLeuPheAlaValLeuCY 859
Db 2788 CTGGGACTGGAGACAGCGCTGCATCCCTTGCAGAAATATGGAGTATTGTCTGTTCTCTG 2847
QY 859 sIleGluThrSerHisTyrValAlaPheValLysTyrGlyLysAspSerAlaTrpLe 879
Db 2848 CATAGAAACAGCCACTATGTTGCTTTGTGAAGTATGGAGGACGATCTGCCTGGCT 2907
QY 879 uPhePheAspSerMetAlaAspArgAspGlyGlnAsnGlyPheAsnIleProGlnVa 899
Db 2908 CTTCTTTGACAGCATGGCCGATCGGATGGTGTGTCAGAAATGGCTTCAACATTCCTCAAGT 2967
QY 899 lThrProCysProGluValGlyGluTyrLeuLysMetSerLeuGluAspLeuHisSerLe 919
Db 2968 CACCCCATGCCAGAAAGTAGGAGAGTACTTGAAGATGTTCTGGAAGACCTGCATTCCTT 3027
QY 919 uAspSerArgArgIleGlnGlyCysAlaArgArgLeuLeuCysAspAlaTyrMetCysMe 939
Db 3028 GGACTCCAGGAGAAATCCAGGCTGTGCAGAAAGCTGTTTGTGATGCATATATGTCAT 3087
QY 939 tTyrGlnSerProThrMetSerLeuTyrLys 949
Db 3088 GTACCAGAGTCCAACTAGTGTGTACAAA 3118
```

## RESULT 4

ABV75394

ID ABV75394 standard; DNA; 5371 BP.

XX AC ABV75394;

XX DT 18-MAR-2003 (first entry)

XX DE Human CYLD polypeptide encoding DNA.

XX KW CD40; NF-kappaB essential modulator; NEMO; CYLD; immunosuppressive;  
XX KW antiinflammatory; cytostatic; gene therapy; human; gene; ds.  
XX OS Homo sapiens.

XX FH Key Location/Qualifiers  
XX FT CDS 392..3262  
XX FT /\*tag= a  
XX FT /product= "CYLD"

XX PN WO200292761-A2.

XX PD 21-NOV-2002.

XX XX 07-MAY-2002; 2002WO-US014570.

XX XX 08-MAY-2001; 2001US-00851673.

XX XX (IMMUNEX ) IMMUNEX CORP.

XX XX Derry JMJ, Fanslow WC, Dougall WC;

XX XX WPI; 2003-120669/11.

XX XX P-PSDB; ABB82783.

XX XX Identifying compounds that alter one or more biological activities of  
XX XX CD40 by modulating the binding of NEMO and CYLD, useful for treating  
XX XX disorders of the immune system, and inflammatory and cancer diseases.  
XX XX Example 5; Page 38-44; 48pp; English.

XX CC The invention relates to methods of identifying compounds that alter one/  
CC more biological activities of CD40. One method involves screening for  
CC molecules that modulate the binding of NF-kappaB essential modulator  
CC (NEMO) and CYLD. The methods and compositions of the invention of  
CC determining compounds that agonize or antagonize a CD40 signaling  
CC activity, are useful for the further definition of CD40-mediated  
CC signaling pathways, and for manipulation of CD40-mediated cellular  
CC responses. They also provide therapeutic agents for treating disorders of  
CC the immune system, and inflammatory and cancer diseases. The present  
CC sequence represents a DNA encoding the human CYLD polypeptide

XX SQ Sequence 5371 BP; 1600 A; 950 C; 1183 G; 1638 T; 0 U; 0 Other;

## Alignment Scores:

Prod. No.: 0 Length: 5371  
Score: 685.00 Matches: 948  
Percent Similarity: 97.63% Conservative: 0  
Best Local Similarity: 97.63% Mismatches: 1  
Query Match: 72.18% Indels: 23  
DB: 10 Gaps: 0

US-09-671-687A-3 (1-949) x ABV75394 (1-5371)

QY 1 MetSerSerGlyLeuTrpSerGlnGluLysValThrSerProTyrTrpGluGluArgile 20  
Db 392 ATGAGTTTCAAGCTTATGGAGCCAAAGAAAGTCACTTCACTTCTGGGAAGAGCGATT 451

QY 21 PheTyrLeuLeuLeuGlnGluCysSerValThrAspLysGlnThrGlnLysLeuLeuLys 40  
Db 452 TTTTACTTGTCTTCAAGAAATGCAGCGTTACAGACAAACAAACACAAAAGCTCTTTAAA 511

QY 41 ValProLysGlySerIleGlyGlnTyrIleGlnAspArgSerValGlyHisSerArgile 60  
Db 512 GTACCGAAGGGAAGTATAGACAGTATATTCAGATCGTGTCTGTGGGGCATTCAGGATT 571

QY 61 ProSerAlaLysGlyLysLysAsnGlnIleGlyLeuLysIleLeuGlnProHisAla 80  
Db 572 CCTTCTGCAAAAGGCAAGAAATAATCAGATTGGATTAAAAATTTAGAGCAACCTCATGCA 631

QY 81 ValLeuPheValAspGlu---AspValValGluIleAsnGluLysPheThrGluLeuLeu 99  
Db 632 GTTCTCTTTGTTGATGAAAGGATGTTGTAGAGATAAATGAAAGTTTACAGAGTTACTT 691

QY 100 LeuAlaIleThrAsnCysGluGluArgPheSerLeuPheLysAsnArgAsnArgLeuSer 119  
Db 692 TTGGCAATTACCAATTTGTGAGGAGAGTTTCAGCTCTTTAAAAACAGAAACAGACTAAGT 751

QY 120 LysGlyLeuGlnIleAspValGlyCysProValLysValGlnLeuArgSerGlyGluGlu 139  
Db 752 AAAGGCTCCAAATAGACGTGGCTGTCTGTGAAAGTACAGCTGAGATCTGGGGAAGAA 811

QY 140 LysPheProGlyValValArgPheArgGlyProLeuLeuAlaGluArgThrValSerGly 159  
Db 812 AAATTTCTGGAGTTGTACGTTTCAGAGACCCCTTGTAGCAGAGAGGACAGTCTCCGA 871

QY 160 IlePhePheGlyValGluLeuLeuGluGlyArgGlyGlnGlyPheThrAspGlyVal 179  
Db 872 ATATTCTTTGGAGTTGAAATTGCTGGAAGAAAGGTCGTGGTCAAGGTTTCTAGTACGGGGTG 931

QY 180 TyrGlnGlyLysGlnLeuPheGlnCysAspGluAspCysGly---PheValAlaLeuAsp 198  
Db 932 TACCAAGGGAACACAGCTTTTTCAGTGTGATGAAGATTGTGGCGTGTGTGTGATCTGGAC 991

QY 199 LysLeuGluLeuIleGluAspAspThrAlaLeuGluSerAspTyrAlaGlyProGly 218  
Db 992 AAGCTAGAACTCATAGAAGATGATGACTGCATTGGAAGATGATTACCGAGGTCTCTGG 1051

QY 219 AspThrMetGlnValGluLeuProProLeuGluIleAsnSerArgValSerLeuLysGly 238  
Db 1052 GACACAATCAGGTGCAACTTCTCTCTTTGGAAATAAATCCACAGAGTTTCTTTTGAAG-GGT 1110

QY 239 -GlyGluThrIleGluSerGlyThrValIlePheCysAspValLeuProGlyLysGluSe 258

[illegible]



XX ADRI4488;  
XX 21-OCT-2004 (first entry)  
XX Human NF-kappaB pathway-associated gene SeqID489.  
XX NF-kappaB pathway; antiinflammatory; cytostatic; hepatotropic; virucide;  
XX antiarthritic; antirheumatic; gastrointestinal-Gen; antiasthmatic;  
XX antiarteriosclerotic; immunomodulator; cerebroprotective; vasotropic;  
XX immunosuppressive; vulnary; gene therapy; immune disorder;  
XX inflammatory disorder; NF-kappaB regulation; cancer; aberrant apoptosis;  
XX hepatic disorder; Hodgkin's lymphoma; haematopoietic tumour;  
XX hyper-IgM syndrome; hypohidrotic ectodermal dysplasia;  
XX X-linked anhidrotic ectodermal dysplasia; immunodeficiency;  
XX viral infection; HIV-1; HTLV-1; hepatitis B; hepatitis C; EBV; influenza;  
XX viral replication; host cell survival; evasion of immune response;  
XX rheumatoid arthritis; inflammatory bowel disease; colitis; asthma;  
XX atherosclerosis; cachexia; euthyroid sick syndrome; stroke; EAE;  
XX autoimmune disorder; hyper immune activity;  
XX aberrant acute phase response; hypercongenital condition; birth defect;  
XX necrotic lesion; wound; organ transplant rejection;  
XX aberrant signal transduction; proliferating disorder; cancer;  
XX HIV propagation; Gene; ds; human.  
XX Homo sapiens.  
XX WO2004065577-A2.  
XX 05-AUG-2004.  
XX 13-JAN-2004; 2004WO-US000798.  
XX 14-JAN-2003; 2003US-0440068P.  
XX 12-MAY-2003; 2003US-0469757P.  
XX (BRIM ) BRISTOL-MYERS SQUIBB CO.  
XX Nadler SG, Neubauer MG, Feder JN, Carman J;  
XX WPI; 2004-562168/54.  
XX P-PSDB; ADRI4489.  
XX New isolated polynucleotides and polypeptides associated with NF-kappaB  
XX pathway, useful for diagnosing, treating, or preventing disorders or  
XX diseases associated with NF-kappaB pathway.  
XX Claim 1; SEQ ID NO 489; 237pp; English.  
XX This invention relates to the novel association of protein sequences (and  
XX the genes which encode them) to the NF-kappaB pathway. The invention may  
XX be useful for the production of compounds with an antiinflammatory,  
XX cytostatic, hepatotropic, virucide, antiarthritic, antirheumatic,  
XX gastrointestinal-Gen, antiasthmatic, antiarteriosclerotic,  
XX immunomodulator, cerebroprotective, vasotropic, immunosuppressive or  
XX vulnary activity or for gene therapy. The proteins and nucleotides are  
XX useful for diagnosing, preventing, treating, or ameliorating conditions  
XX or diseases associated with the NF-kappaB pathway. The condition is an  
XX immune disorder, an inflammatory disorder, an inflammatory disorder  
XX related to aberrant NF-kappaB regulation, cancer, aberrant apoptosis,  
XX hepatic disorders, Hodgkin's lymphomas, haematopoietic tumours, hyper-IgM  
XX syndromes, hypohidrotic ectodermal dysplasia, X-linked anhidrotic  
XX ectodermal dysplasia, immunodeficiency, viral infections, HIV-1, HTLV-1,  
XX hepatitis B, hepatitis C, EBV, influenza, viral replication, host cell  
XX survival, evasion of immune responses, rheumatoid arthritis, inflammatory  
XX bowel disease, colitis, asthma, atherosclerosis, cachexia, euthyroid sick  
XX syndrome, stroke, EAE, autoimmune disorders, disorders related to hyper  
XX immune activity, disorders related to aberrant acute phase responses,  
XX hypercongenital conditions, birth defects, necrotic lesions, wounds,  
XX organ transplant rejection, conditions related to organ transplant  
XX rejection, disorders related to aberrant signal transduction,  
XX proliferating disorders, cancers and HIV propagation in cells infected  
XX with other viruses. The present sequence is that of a human gene which is

CC subject to the novel association with the NF-kappaB pathway of the  
CC invention. Note: This sequence does not appear in the specification but  
CC was obtained by the indexer from Genbank.  
XX  
SQ Sequence 5371 BP; 1600 A; 950 C; 1183 G; 1638 T; 0 U; 0 Other;  
Alignment Scores:  
Pred. No.: 0 Length: 5371  
Score: 685.00 Matches: 948  
Percent Similarity: 97.63% Conservative: 0  
Best Local Similarity: 97.63% Mismatches: 1  
Query Match: 123 Indels: 23  
DB: 13 Gaps: 0  
US-09-671-687A-3 (1-949) x ADRI4488 (1-5371)  
QY 1 MetSerSerGlyLeuTrpSerGlnGluLysValThrSerProTyrTrpGluGluArgile 20  
Db 392 ATGAGTTTCAGGCTTATGAGGCCAAGAAAGTCACTTCACTTCTGCGAAGAGCGGATT 451  
QY 21 PheTyrLeuLeuGlnGluCysSerValThrAspLysGlnThrGlnLysLeuLys 40  
Db 452 TTTTACTTGTCTTCTTCAAGAATCGAGCTTACAGACAAACACAAAAAGCTCTTTAAA 511  
QY 41 ValProLysGlySerIleGlyGlnTyrIleGlnAspArgSerValGlyHisSerArgile 60  
Db 512 GTACCGAGGGAAGTATAGGACAGTATATTCAGATCGTCTGTGGGCAATTCAGAGATT 571  
QY 61 ProSerAlaLysGlyLysLysAsnGlnIleGlyLeuLysIleLeuGlnProHisAla 80  
Db 572 CCTTCTGCAAAAGGCAAGAAATCAGATTGGATTAAATAATCTAGAGCAACCTCATGCA 631  
QY 81 ValLeuPheValAspGlu---AspValValGluLeuAsnGlnLysPheThrGluLeuLeu 99  
Db 632 GTTCTCTTTGTGATGAAAGGATGTTGTAGAGATAAATGAAAAGTTTCAGAGATTACTT 691  
QY 100 LeuAlaIleThrAsnCysGluGluArgPheSerLeuPheLysAsnArgAsnArgLeuSer 119  
Db 692 TTGGCAATTACCAATTTGTGAGGAGAGTTTCAGCTGTTTAAAAACAGAAACAGACTAAGT 751  
QY 120 LysGlyLeuGlnIleAspValGlyCysProValLysValGlnLeuArgSerGlyGluGlu 139  
Db 752 AAAGGCTCCCAATACAGCTGGGCTGTCTGTGAAAGTACAGCTGAGATCTGGGGAAGAA 811  
QY 140 LysPheProGlyValValArgPheArgGlyProLeuLeuAlaGluArgThrValSerGly 159  
Db 812 AAATTTCTCGAGTTGTACGCTTTCAGAGGACCCCTGTTAGCAGAGAGACAGTCTCCGGA 871  
QY 160 IlePhePheGlyValGluLeuLeuGluGlyArgGlyGlnGlyPheThrAspGlyVal 179  
Db 872 ATATTCTTTGGAGTTGAAATTTGCTGGGAAGAGGTCGTGGTCAAGGTTTCACTGACGGGGTG 931  
QY 180 TyrGlnGlyLysGlnLeuPheGlnCysAspGluAspCysGly---PheValAlaLeuAsp 198  
Db 932 TACCAGGGAACAGCTTTTTCAGTGTGATGAGATTGTGGCGTGTGTTGTGATGGAC 991  
QY 199 LysLeuGluLeuLeuGluAspAspThrAlaLeuGluSerAspTyrAlaGlyProGly 218  
Db 992 AAGCTAGAACTCATAGACATGATGACATGCTGCTTGGAAAGTGAATGATTCAGGCTCTGGG 1051  
QY 219 AspThrMetGlnValGluLeuProLeuGluLeuLeuAsnSerArgValSerLeuLysGly 238  
Db 1052 GACACAATGAGTGTGAACTCTCTCTTTTGGAAATAAATCACTCCAGAGTTTCTTTGAA--GGT 1110  
QY 239 -GlyGluThrIleGluSerGlyThrValIlePheCysAspValLeuProGlyLysGluSe 258  
Db 1111 TGGAGAAACATAGAAATCTGGAACAGTTATATCTGTGATGTTTGGCCAGAAAAGAAAG 1170  
QY 258 rLeuGlyTyrPheValGlyValAspMetAspAsnProIleGlyAsnTrpAspGlyArgph 278  
Db 1171 CTTAGGATATTTTGTGTTGGTGTGACATGGATAACCTATTGGCAACTGGGATGGAAGATT 1230  
QY 278 eAspGlyVal---LeuCysSerPheAlaCysValGluSerThrIleLeuLeuHisIleAs 297



1231 TGATGAGTGCAGCTTTGAGTTTTGCGTGTGTGAAGTACAAATCTATTGCAATCAA 1290  
297 nAspIlelePro-----GluSerValThrGlnGluArgArgProProLysLeuAl 314  
1291 TGATATATCCAGCTTATATCAGAGAGTGTGACGCGAGGAAGAGGCGCTCCCAACTTGC 1350  
314 aPheMetSerArgGlyValGlyAspLysGlySerSerHisAsnLysProLysAlaTh 334  
1351 CTTTATGTCAAGAGGTGTGGGCAAAAGGTTTCATCCAGTCATAATAAACCAAGGCTAC 1410  
334 rGlySerThrSerAspProGlyAsnArg---ArgSerGluLeuPheTyrThrLeuAsnGl 353  
1411 AGGATCTACCTCAGACCCCTGGAAATAGAAACAGATCTGAATATTATTTATACCTTAAATGG 1470  
353 ySerSerValAspSerGlnProGlnSerLysSerLysAsnThrTTPtTyrIleAspGluVa 373  
1471 GTCTTCTGTGACTCACAACCAATCCAAATCAAAAATACATGGTACATTGATGAAGT 1530  
373 lAlaGluAspProAlaLysSerLeuThrGluLeuSerThrAspPheAspArgSerPr 393  
1531 TGCAGAAGACCCCTGCAAAATCTTTACAGAGATATCTACAGACTTTGACCGTTCTTACC 1590  
393 oProLeuGlnProProProValAsnSerLeuThrThrGluAsnArgPheHisSerLeuPr 413  
1591 ACCACTCAGCCCTCTCTGTAATCTCACTGACCCAGAGAGTTCAGACTTTTACCTCTTTACC 1650  
413 oPheSerLeuThrLysMetProAsnThrAsnGlySerIleGlyHisSerProLeuSerLe 433  
1651 ATTCACTCTCACCAGAGTCCCAATACCAATGGAAGTATTGGCCACAGTCCACTTCTCT 1710  
433 uSerAlaGlnSerValMetGluLeuAsnThrAlaProValGlnGluSerProProle 453  
1711 GTCAGCCCACTGTGTAATGGAAGAGCTAAACACTGCACCCGTCGAAGAGAGTCCACCTT 1770  
453 uAlaMetProProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLysGl 473  
1771 GGCCATGCTCTCTGGGAATCACTATGGTCTAGAAGTGGGCTCATTTGGCTGAAGTTAAGA 1830  
473 uAsnProProPheTyrGlyValIleArgTrpIleGlyGlnProProGlyLeuAsnGluVa 493  
1831 GAACCTCTCTTATATGGGTNANTCCGTTGGATCGGTACGCCACAGGACTGAATGAAGT 1890  
493 lLeuAlaGlyLeuGluLeuGluAspGluCysAlaGlyCysThrAspGlyThrPheArgGl 513  
1891 GCTGCTGGACTGGAACCTGGAAGATGAGTGTGACGCTGTACGGATGGAACTTCAGAGG 1950  
513 yThrArgTyrPheThrCysAlaLeuLysLysAlaLeuPheValLysLeuLysSerCysAr 533  
1951 CACTCGGTATTTCACCTGTGCCCTGAAGAAGCGCTGTTTGTGAACCTGAAGAGCTGCAG 2010  
533 gProAspSerArgPheAlaSerLeuGlnProValSerAsnGlnIleGluArgCysAsnSe 553  
2011 GCCGTACTCTAGTTTGTATCATCTTCAGCCGGTTTCCAATCAGATTGAGCGCTGTAACCT 2070  
553 rLeuAlaPheGlyGlyTyrLeuSerGluValValGluGluAsnThrProProLysMetGl 573  
2071 TTTAGCATTTGGAGGCTACTTAAGTAGAGTAGTAGAAGAAATATCTCCACCAAAATGGA 2130  
573 uLysGluGlyLeuGluIleMetIleGlyLysLysGlyIleGlnGlyHisTyrAsnSe 593  
2131 AAAAGAAGGCTGGAGATAATGATTGGGAAGAAAGAGCATCCAGGCTCATTAACAATTC 2190  
593 rCysTyrLeuAspSerThrLeuPheCysLeuPheAlaPheSerSerValLeuAspThrVa 613  
2191 TTGTTACTTAGACTCAACCTTATCTGCTTATTTCTTTAGTCTTGTCTGTGACACTGT 2250  
613 lLeuLeuArgProLysGlyLysAsnAspValGluTyrTyrSerGluThrGlnGluLeuLe 633  
2251 GTTACTTAGACCAAGAAAGAACGATGTAGANATATTATAGTGAACCCAGAGACTACT 2310  
633 uArgThrGluIleValAsnProLeuArgIleTyrGlyTyrValCysAlaThrLysIleMe 653

2311 GAGGACAGAAATGTTAATCTCTCGAATATATATCGAATATATGTGTGCCACAAAATTAT 2370  
653 tLysLeuArgLysIleLeuGluLysValGluAlaAlaSerGlyPheThrSerGluGluLy 673  
2371 GAAACTGAGGAAATACTTTGAAAAGGTGGAGGCTGCATCAGATTTACTCTGAAGAAA 2430  
673 sAspProGluGluPheLeuAsnIleLeuPheHisIleLeuArgValGluProLeuLe 693  
2431 AGATCCTCAGGAATCTTGAATATTCTGTTTCATCATATTTTAAGGGTAGAACCTTGTCT 2490  
693 uLysIleArgSerAlaGlyGlnLysValIleAspCysTyrPheTyrGlnIlePheMetGl 713  
2491 AAAAATAAGATCAGCAGGTCAAAAGGTACAAGATTGTACTTCTATCAAAATTTTATGGA 2550  
713 uLysAsnGluLysValGlyValProThrIleGlnGlnLeuLeuGluTTPSerPheIleAs 733  
2551 AAAAATAGAAAGTTGGCGTTCACCAATTCAGCAGTGTGTAGAAATGGTCTTTTATCA 2610  
733 nSerAsnLeuLysPheAlaGluAlaProSerCysLeuIleIleGlnMetProArgPheGl 753  
2611 CAGTAACCTGAAATTTGCAGAGGCACCATCATGTCTGATTATTACAGTGCCTCGATTGG 2670  
753 yLysAspPheLysLeuPheLysLysIlePheProSerLeuGluLeuAsnIleThrAspLe 773  
2671 AAAAGACTTTAAACTATTATAAAAAATTTTCTCTCTGGAATTTAAATATAACAGATT 2730  
773 uLeuGluAspThrProArgGlnCysArgIleCysGlyLysLeuAlaMetTyrGluCysAr 793  
2731 ACTTGAAGACACTCCACAGACAGTCCGGATATGTGGAGGCTTGCANATGATAGTGTAG 2790  
793 gGluCysTyrAspAspProAspIleSerAlaGlyLysIleLysGlnPheCysLysThrCy 813  
2791 AGAATGCTACGACGATCCGGACATCTCAGCTGGAAAAATCAAGCAGTTTGTAAAACTG 2850  
813 sAsnThrGlnValHisLeuHisProLysArgLeuAsnHisLysTyrAsnProValSerLe 833  
2851 CAACACTCAAGTCCACCTTCATCCAGAGGCTGAATATAAATAAACCAGTGTCTACT 2910  
833 uProLysAspLeuProAspTTPAspTTPArgHisGlyCysIleProCysGlnAsnMetGl 853  
2911 TCCCAAGACTTACCCGACTGGGACTGGAGACACGGCTGCATCCCTTCCAGAAATATGGA 2970  
853 uLeuPheAlaValLeuCysIleGluThrSerHisTyrValAlaPheValLysTyrGlyLy 873  
2971 GTTATTTCGTCTCTCATAGAAACCAAGCCACTATCTTCTTTTGTGAAGTATGGGAA 3030  
873 sAspAspSerAlaTTPLeuPhePheAspSerMetAlaAspArgAspGlyGlyGlnAsnGl 893  
3031 GGACGATTCTGCTGGCTCTTCTTTGACAGCATGGCCGATCGGGATGGTGTGAGATGG 3090  
893 yPheAsnIleProGlnValThrProCysProGluValGlyGluTyrLeuLysMetSerLe 913  
3091 CTTCAACATTCCTCAAGTCAACCCATGCCCAAGATAGAGAGTACTTTGAAGATGTCTCT 3150  
913 uGluAspLeuHisSerLeuAspSerArgIleGlnGlyCysAlaArgArgLeuLeuCy 933  
3151 GGAAGACCTGATCTCTTGGACTCCAGAGAAATCCAGGCTGTGACCAAGACTCTTTG 3210  
933 sAspAlaTyrMetCysMetTyrGlnSerProThrMetSerLeuTyrLys 949  
3211 TGATGCATATATGTGCATGTACCAGAGTCCAAACAATGAGTTTGTACAAA 3259  
RESULT 6  
ADQ95915  
ID ADQ95915 standard; cDNA; 3302 BP.  
XX  
AC ADQ95915;  
XX  
DT 07-OCT-2004 (first entry)  
XX  
DE T cell activation associated cDNA #47.  
XX  
KW ss; gene; antiallergic; antiarthritic; antiasthmatic; antidiabetic;

KW anti-HRV; antimicrobial; antirheumatic; immunosuppressive;  
 KW neuroprotective; gene therapy; T cell activation; diagnosis;  
 KW autoimmune disease; rheumatoid arthritis; asthma; multiple sclerosis;  
 KW diabetes; allergic disease; infectious disease; AIDS; chronic rejection;  
 KW organ; bone-marrow transplant.

XX Homo sapiens.

XX Key Location/Qualifiers  
 FT CDS 243..3104  
 FT /\*tag= a

XX WO2004058805-A2.

XX 15-JUL-2004.

XX 25-DEC-2003; 2003WO-JP016715.

XX 26-DEC-2002; 2002JP-00376365.

XX 27-DEC-2002; 2002US-0436473P.

XX 25-APR-2003; 2003JP-00122113.

XX 28-APR-2003; 2003US-0465792P.

XX 21-OCT-2003; 2003JP-00360559.

XX 22-OCT-2003; 2003US-0512846P.

XX (ASAH-) ASahi KASEI PHARMA CORP.

XX Matsuda A, Yoneta S;

XX WPI; 2004-593134/57.

XX P-PSDB; ADQ95916.

XX New purified protein involved in T cell activation, useful for  
 PT diagnosing, preventing and/or treating acquired immunodeficiency  
 PT syndrome, autoimmune (e.g. rheumatoid arthritis, and diabetes), allergic  
 PT and infectious diseases.

XX Claim 4; SEQ ID NO 93; 2828pp; English.

XX The invention relates to purified proteins and genes encoding them, that  
 CC are involved in T cell activation (I) and has an amino acid deletion,  
 CC substitution or addition in the amino acid sequences. The methods and  
 CC compositions of the present invention are useful for the diagnosis,  
 CC prevention and/or treatment of autoimmune disease (rheumatoid arthritis,  
 CC asthma, multiple sclerosis and diabetes), allergic disease, infectious  
 CC disease, AIDS, and acute or chronic rejection at organ transplant or bone  
 CC -marrow transplant. This sequence corresponds to a cDNA for a protein  
 CC involved in T cell activation.

XX SQ Sequence 3302 BP; 981 A; 648 C; 772 G; 901 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 0 Length: 3302  
 Score: 667.00 Matches: 947  
 Percent Similarity: 98.34% Conservative: 0  
 Best Local Similarity: 98.34% Mismatches: 2  
 Query Match: 70.28% Indels: 16  
 DB: 12 Gaps: 0

US-09-671-687A-3 (1-949) x ADQ95915 (1-3302)

QY 1 MetSerSerGlyLeuThrSerGlnGluLysValThrSerProTyrTrpGluLysArgIle 20  
 Db 243 ATGAGTTCAGGCTTATGAGGCAAGAAAGTCACTTACCTACTGGGAAGCGGATT 302

QY 21 PheTyrLeuLeuGlnGluCysSerValThrAspLysGlnThrGlnLysLeuLeuLys 40  
 Db 303 TTTTACTTGCTTCTCAGAGATGACGGTTACAGACAAACAAACAAAGTCTCTTAA 362

QY 41 ValProLysGlySerIleGlnTyrIleGlnAspArgSerValGlyHisSerArgIle 60  
 Db 363 GTACCGAAGGAAGTATAGGACAGTATATTCAAGATCGTTCGTGGGGCATTCAAGGATT 422

QY 61 ProSerAlaLysGlyLysLysAsnGlnIleGlyLeuLysIleLeuGluGlnProHisAla 80  
 Db 423 CTTTCTGCAAAAGGCAAGAAATCAGATTGGATTAAAAATTTAGAGCAACCTCATGCA 482  
 QY 81 ValLeuPheValAspGlu---AspValValGluIleAsnGluLysPheThrGluLeuLeu 99  
 Db 483 GTTCTCTTTGTCATGAAAAGGATGTTGTAGAGATAAATGAAAGTTTCAGAGTTACTT 542  
 QY 100 LeuAlaIleThrAsnCysGluGluArgPheSerLeuPheLysAsnArgAsnArgLeuSer 119  
 Db 543 TTGGCAATTACCAATTGTGAGGAGAGGTTGAGCTCTGTGAAGTACAGCTGAGATCTGGGAAGAA 602  
 QY 120 LysGlyLeuGlnIleAspValGlyCysProValLysValGlnLeuArgSerGlyGluGlu 139  
 Db 603 AAAGGCTCCAAATAGACGTTGGCTGTCTGTGAAGTACAGCTGAGATCTGGGAAGAA 662  
 QY 140 LysPheProGlyValValArgPheArgGlyProLeuLeuAlaGluArgThrValSerGly 159  
 Db 663 AAATTTCTGGAGTTGTACGCTTCAGAGGACCCCTGTAGCAGAGAGGACAGTCTCCGGA 722  
 QY 160 IlePhePheGlyValGluLeuLeuGluGlyArgGlyGlnGlyPheThrAspGlyVal 179  
 Db 723 ATATTTCTTTGGAGTTGAATTGCTGGAAGAGGTCGTGGTCAAGGTTTCACTGACGGGGT 782  
 QY 180 TyrGlnGlyLysGlnLeuPheGlnCysAspGluAspCysGly---PheValAlaLeuAsp 198  
 Db 783 TACCAAGGGAACAGCTTTTTCAGTGTGATGAAGATTGTGGCTGTTTGTTCATTGGAC 842  
 QY 199 LysLeuGluLeuIleGluAspAspThrAlaLeuGluSerAspTyrAlaGlyProGly 218  
 Db 843 AAGCTAGAACTCATAGAGATGATGACACTGCATTGGAAGTGAATGACGAGGCTCTCTGGG 902  
 QY 219 AspThrMetGlnValGluLeuProLeuGluLysAsnSerArgValSerLysGly 238  
 Db 903 GACACAATCAGGTCGAACCTCTCTTTGGAATAAATCCAGAGTTTCTTTGAA--GGT 961  
 QY 239 -GlyGluThrIleGluSerGlyThrValIlePheCysAspValLeuProGlyLysGluSe 258  
 Db 962 TGGAGAAACAATAGAAATCTGGAACAGTTATATCTGTGATGTTTTCAGGAAAGAAAG 1021  
 QY 258 rLeuGlyTyrPheValGlyValAspMetAspAsnProIleGlyAsnTrpAspGlyArgPh 278  
 Db 1022 CTTAGGATATTTTGTGGTGTGACATGATAACCTATTTGGCACTGGGATGGAAGATT 1081  
 QY 278 eAspGlyVal---LeuCysSerPheAlaCysValGluSerThrIleLeuLeuHisIleAs 297  
 Db 1082 TGATGGAGTGCAGCTTTGTAGTTTTCGTGTGTGTAAGATACAATTTCTATTGCACATCAA 1141  
 QY 297 nAspIleIleProGluSerValThrGlnGluArgArgProProLysLeuAlaPheMetSe 317  
 Db 1142 TGATATCATCCAGAGAGTGTGACGAGGAAGAGGCGCTCCCAAACTTGCTTTATGTC 1201  
 QY 317 rArgGlyValGlyAspLysGlySerSerHisAsnLysProLysAlaThrGlySerTh 337  
 Db 1202 AAGAGGTGTGGGCAAGAGTTTCATCCAGTCATATATAACCAAGGCTACAGATCTAC 1261  
 QY 337 rSerAspProGlyAsnArg---ArgSerGluLeuPheTyrThrLeuAsnGlySerSerVa 356  
 Db 1262 CTGAGACCTCGGAAATAGAAACAGATCTGAATTTATTTTATACCTTAAATGGGCTCTCTGT 1321  
 QY 356 lAspSerGlnProGlnSerLysSerLysAsnThrTrpTyrIleAspGluValAlaGluAs 376  
 Db 1322 TGACTCACCAACCAATCCAAATCAAAAATACATGTTGATGTAAGTTGCAGANGA 1381  
 QY 376 pProAlaLysSerLeuThrGluIleSerThrAspPheAspArgSerSerProProLeuGlu 396  
 Db 1382 CCCTGCAAAATCTCTTACAGAGATATCTACAGACTTTGACCGTCTTCCACCACCTCCA 1441  
 QY 396 nProProProValAsnSerLeuThrThrGluAsnArgPheHisSerLeuProPheSerLe 416  
 Db 1442 GCCTCTCTCTGTGAATCACTGACCAACGAGAACAGATTCACACTCTTTTACATTCAGTCT 1501  
 QY 416 uThrLysMetProAsnThrAsnGlySerIleGlyHisSerProLeuSerLeuSerAlaGl 436

Db	1502		CACCAAGATGCCCAATACCAATGGAAGTATTGGCCACAGTCCACTTTCTCTGTGACGCCCA	1561
Qy	436	nSerValMetGluGluLeuAsnThrAlaProValGlnGluSerProProLeuAlaMetPr	456	
Db	1562	GTCTGTAAATGGAAGAGCTAAACACTGCACCGCTCAGAGAGTCCACCCCTTGCCATGCC	1621	
Qy	456	oProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLysGluAsnProPr	476	
Db	1622	TCCTGGGAACCTCACATGGCTAGAAAGTGGGCTCATTTGGCTGAAGTTAAGGAGAACCCCTCC	1681	
Qy	476	oPheTyrGlyValIleArgTrpIleGlyGlnProProGlyLeuAsnGluValLeuAlaGl	496	
Db	1682	TTTCTATGGGGTAATCCGTTGGATGGGTGAGCCACCGAGACTGAATGAAGTGTCTCGCTGG	1741	
Qy	496	YLeuGluLeuGluAspGluCysAlaGlyCysThrAspGlyThrPheArgGlyThrArgTyr	516	
Db	1742	ACTGGAACCTGGAAGATGAGTGTGCAGGCTGTACGATGGAACTTCAGAGGCACCTCGGTA	1801	
Qy	516	rPheThrCysAlaLeuLysLysAlaLeuPheValLysLeuLysSerCysArgProAspSe	536	
Db	1802	TTTCACCTGTGCCCTGAAGAAGGGCGCTGTTGTGAAACTGAGAGCTGCAGGCTGACTC	1861	
Qy	536	xArgPheAlaSerLeuGlnProValSerAsnGlnIleGluArgCysAsnSerLeuAlaPh	556	
Db	1862	TAGGTTTGGCATCATTTGCAGCGGTTTCCATCATCAGATTGAGCGCTGTAACTCTTTAGCATT	1921	
Qy	556	eGlyGlyTyrLeuSerGluValValGluGluAsnThrProProLysMetGluLysGluGl	576	
Db	1922	TGGAGGCTACTTAAGTGAAGTAGTAGAAGAAAATACTCCACAAAATGGAANAAGANG	1981	
Qy	576	YLeuGluIleMetIleGlyLysLysLysGlyIleGlnGlyHisTyrAsnSerCysTyrLe	596	
Db	1982	CTTGAGATAATGATTGGGAAGAAGAAAGGCATCCAGGGTCATTACAAATCTCTGTACTT	2041	
Qy	596	uAspSerThrLeuPheCysLeuPheAlaPheSerSerValLeuAspThrValLeuLeuAr	616	
Db	2042	AGACTCAACCTATTCTGCTATTATTGCTTTTGTGTTCTGTCTGGACACTGTGTACTTAG	2101	
Qy	616	gProLysGluLysAsnAspValGluTyrTyrSerGluThrGlnGluLeuLeuArgThrGl	636	
Db	2102	ACCCANAAGAAAAGAACGATGTAGAAATATTATAGTGAACCCAGAGAGCTCTGAGGACAGA	2161	
Qy	636	uIleValAsnProLeuArgIleTyrGlyTyrValCysAlaThrLysIleMetLysLeuAr	656	
Db	2162	AAATGTTAACTCTGAGATATATGATGATGTTGTGTGCCACAAAAATTTATGAACTGAA	2221	
Qy	656	g-LysIleLeuGluLysValGluAlaIleAspGlyPheThrSerGluGluLysAspProG	676	
Db	2222	-GAAAATACCTGAAAAGGTGGAGCTGCATCAGGATTTAGCTCTCGAAGAAAAAGATCCGT	2280	
Qy	676	luGluPheLeuAsnIleLeuPheHisIleLeuArgValGluProLeuLeuLysIleA	696	
Db	2281	AGGAATTTCTGAAATATTCTGTTTCATCATATTTTAAAGGTAGAACCTTTCTCTAAAATAA	2340	
Qy	696	rgSerAlaGlyGlnLysValGlnAspCysTyrPheTyrGlnIlePheMetGluLysAsnG	716	
Db	2341	GATCAGCAGGTCAAAAGGTACAAGATTTGTTACTTCTATCAAAATTTTATGGAAGAAAAATG	2400	
Qy	716	luLysValGlyValProThrIleGlnLeuLeuGluTrpSerPheIleAsnSerAsnL	736	
Db	2401	AGAAAGTTGGCGTTCCACAAATTCAGAGTTGTTAGAAATGCTCTTTTATCAACAGATAACC	2460	
Qy	736	eLysPheAlaGluAlaProSerCysLeuIleIleGlnMetProArgPheGlyLysAspP	756	
Db	2461	TGAAATTTGCAGAGCCACCATCAGTCTGATTATTTCAGATGCCCTCGATTTTGSAAAAGACT	2520	
Qy	756	heLysLeuPheLysLysIlePheProSerLeuGluLeuAsnIleThrAspLeuLeuAla	776	
Db	2521	TTAAACTATTAAAAAATTTTCTCTCTGGAAATTAATATATACAGATTTACTTGAAG	2580	
Qy	776	spThrProArgGlnCysArgIleCysGlyGlyLeuAlaMetTyrGluCysArgGluCyst	796	

Db	2581	ACACTCCACAGCAGTGC	CGGATATGTG	CGAGCGCTTGC	AAATGTATGTAGTGTAGAGAATGCT	2648
Qy	796	yrAspAspProAspIleSerAlaGlyLysIleLysGlnPheCysLysThrCysAsnThrG	816			
Db	2641	ACGACGATCGGACATCTCAGCTGGAAAAATCAAGCAGTTTGTAAACCTGCACACTC	2700			
Qy	816	lnValHisLeuHisProLysArgLeuAsnHisLysTyrAsnProValSerLeuProLysA	836			
Db	2701	AAGTCCACCTTCATCCGAGAGGCTGAATCATAAATATAACCCAGTGTCACTTCCCAAAG	2760			
Qy	836	spLeuProAspTrpAspTrpArgHisGlyCysIlePheCysGlnAsnMetGluLeuPheA	856			
Db	2761	ACTTACCCGACTGGGACTGGAGACACGCGTGCATCCCTTCCCAAGATATGGAGTTATTG	2820			
Qy	856	laValLeuCysIleGluThrSerHisTyrValAlaPheValLysTyrGlyLysAspAsps	876			
Db	2821	CTGTTCCTTCATAGAAACAAAGCCACTATGTTGCTTTTGTGAAGTATGGGAAGGACGATT	2880			
Qy	876	erAlaTrpLeuPheAspSerMetAlaAspArgAspGlyGlnAsnGlyPheAsnI	896			
Db	2881	CTGCCTGCCTCTCTTTGACAGCATGCCGATCGGATGCTGTCAGAAATGGCTTCAACA	2940			
Qy	896	leProGlnValThrProCysProGluValGlyGluTyrLeuLysMetSerLeuGluAspL	916			
Db	2941	TTCCTCAAGTCACCCCATGCCCAAGTAGGAGAGTACTTGAAGATGTTCTCTGGAAGACC	3000			
Qy	916	euHisSerLeuAspSerArgArgIleGlnGlyCysAlaArgArgLeuLeuCysAspAlat	936			
Db	3001	TGCATTCTTTGGACTCCAGAGAAATCCAGGCTGTGCACGAGACGTCTTTGTGATGCAT	3060			
Qy	936	yrMetCysMetTyrGlnSerProThrMetSerLeuTyrLys	949			
Db	3061	ATATGTGCTGTACACAGTCCCAACCAATGAGTTTGTACAAA	3101			
RESULT 7						
AAI58410						
ID	AAI58410 standard; cdna; 2523 BP.					
XX	AAI58410;					
XX	22-OCT-2001 (first entry)					
DE	Human polynucleotide SEQ ID NO 613.					
KW	Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;					
KW	peripheral nervous system; neuropathy; central nervous system; CNS;					
KW	Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;					
KW	amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;					
KW	chemokinetic; thrombolytic; drug screening; arthritis; inflammation;					
XX	leukaemia; ss.					
OS	Homo sapiens.					
XX	WO200153312-A1.					
PD	26-JUL-2001.					
XX	26-DEC-2000; 2000WO-US034263.					
XX	23-DEC-1999; 99US-00471275.					
PR	21-JAN-2000; 2000US-0048725.					
PR	25-APR-2000; 2000US-0052317.					
PR	20-JUN-2000; 2000US-00598042.					
PR	19-JUL-2000; 2000US-00620312.					
PR	03-AUG-2000; 2000US-00653450.					
PR	14-SEP-2000; 2000US-00662191.					
PR	19-OCT-2000; 2000US-00693036.					
PR	29-NOV-2000; 2000US-00727344.					
PA	(HYSE-) HYSEQ INC.					
XX	Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;					
PI	Wang JY, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao Q;					

PI Zhou P, Goodrich R, Drmanac RT;  
XX WPI; 2001-442253/47.  
DR P-PSDB; AAM39254.  
XX Novel nucleic acids and polypeptides, useful for treating disorders such  
PT as central nervous system injuries.  
XX Claim 1; SEQ ID NO 613; 10078pp; English.  
XX The invention relates to human nucleic acids (AAI57798-AAI61369) and the  
CC encoded polypeptides (AAM38642-AAM42213) with nootropic,  
CC immunosuppressant and cytostatic activity. The polynucleotides are useful  
CC in gene therapy. A composition containing a polypeptide or polynucleotide  
CC of the invention may be used to treat diseases of the peripheral nervous  
CC system, such as peripheral nervous injuries, peripheral neuropathy and  
CC localised neuropathies and central nervous system diseases, such as  
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
CC utilisation of the activities such as: Immune system suppression,  
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
CC assays for receptor activity, arthritis and inflammation, leukaemia and  
CC C.N.S disorders. Note: The sequence data for this patent did not form  
CC part of the printed specification  
XX Sequence 2523 BP; 757 A; 503 C; 557 G; 706 T; 0 U; 0 Other;  
Alignment Scores:  
Pred. No.: 0 Length: 2523  
Score: 641.00 Matches: 668  
Percent Similarity: 99.55% Conservative: 0  
Best Local Similarity: 99.55% Mismatches: 0  
Query Match: 67.54% Indels: 3  
DB: 4 Gaps: 0  
US-09-671-687A-3 (1-949) x AAI58410 (1-2523)  
QY 282 LeuCysSerPheAlaCysValGluSerThrIleLeuLeuHisIleAsnAspIlelePro 301  
DB 129 CTTTGTAAGTTGGCTGTGTTGAAAGTACAAATTCATTGGACATCAATGATATCATCCCA 188  
QY 302 GluSerValThrGlnGluArgProProlysLeuAlaPheMetSerArgGlyValGly 321  
DB 189 GAGAGTGTGCGCAGGAAGAGGCTCCCAACTTGTCTTATGTCAAGAGGTGTGG 248  
QY 322 AspLysGlySerSerHisAsnLysProlysAlaThrGlySerThrSerAspProGly 341  
DB 249 GACAAAGGTTCCAGTCATATAATAAACCAAGGCTACAGGATCTACCTCAGACCTGGA 308  
QY 342 AsnArg---ArgSerGluLeuPheThrThrLeuAsnGlySerSerValAspSerGlnPro 360  
DB 309 AATAGAAACAGATCTGAATATTTATACCTTAAATGGGTCTCTGTGTGACTCACAAACA 368  
QY 361 GlnSerLysSerLysAsnThrTrpThrIleAspGluValAlaGluAspProAlaLysSer 380  
DB 369 CAATCCAAATCAAAAAATACATGGTACATGATGAAGTTGACAGAGACCCCTGCAAAATCT 428  
QY 381 LeuThrGluIleSerThrAspPheArgSerSerProProLeuGlnProProVal 400  
DB 429 CTTACAGAGATATCTACAGACTTTGACCGTCTTCCACCACTCTCAGCCTCTCTGTG 488  
QY 401 AsnSerLeuThrThrGluAsnArgPheHisSerLeuProPheSerLeuThrLysMetPro 420  
DB 489 AACTCACTGACCCAGAGACAGATTCACCTCTTTTACCATTCCAGTCTCACCAAGATGCC 548  
QY 421 AsnThrAsnGlySerIleGlyHisSerProLeuSerLeuSerAlaGlnSerValMetGlu 440  
DB 549 AATACCAGTGAAGTATTGGCCACAGTCCACTTTCTCTGTACGCCAGCTCTGTAATGGA 608  
QY 441 GluLeuAsnThrAlaProValGlnGluSerProProLeuAlaMetProProGlyAsnSer 460  
DB 609 GAGCTAACACTGCACCCGTCACAGAGAGTCCACCTTGGCCATGCTCTCTGGGAACTCA 668

QY 461 HisGlyLeuGluValGlySerLeuAlaGluValLysGluAsnProProPheThrGlyVal 480  
DB 669 CATGGCTAGAAAGTGGGCTCATTTGGCTGAAGTTAAGAGAACCCCTCTCTTATGGGTA 728  
QY 481 IleArgTrpIleGlyGlnProProGlyLeuAsnGluValLeuAlaGlyLeuGluGlu 500  
DB 729 ATCCGTTGGATCGGTGAGCCAGGACCTGAAGTGTCTGCTGAGCTGGAACTGGAA 788  
QY 501 AspGluCysAlaGlyCysThrAspGlyThrPheArgGlyThrArgThrPheThrCysAla 520  
DB 789 GATGAGTGTGCGAGGCTGTACGATGGAACCTTCAGAGGCACTCGGTATTTTCACTGTGCC 848  
QY 521 LeuLysLysAlaLeuPheValLysLysSerCysArgProAspSerArgPheAlaSer 540  
DB 849 CTGAAGAGGCGCTGTTGTGAACTGAAGAGCTGAGGCTGACTCTAGGTTCATCA 908  
QY 541 LeuGlnProValSerAsnGlnIleGluArgCysAsnSerLeuAlaPheGlyGlyThrLeu 560  
DB 909 TTTCAGCGGTTTCCAAATCAGATTGAGCGCTGTAACTCTTTAGCAATTTGGAGGCTACTTA 968  
QY 561 SerGluValValGluGluAsnThrProProLysMetGluLysGluGluLysMet 580  
DB 969 AGTGAAGTAGTAGAAGAAATATCTCCACCAAAATGGAAGAGAGGCTTGGAGATAATG 1028  
QY 581 IleGlyLysLysGlyIleGlnGlyHisThrAsnSerCysThrLeuAspSerThrLeu 600  
DB 1029 ATTGGAGAGAGAGGCGATCCAGGCTCAATTCATTTCTTTTACTTAGACTCAACTTA 1088  
QY 601 PheCysLeuPheAlaPheSerValLeuAspThrValLeuLeuArgProLysGluLys 620  
DB 1089 TTCTGCTTATTTGCTTTTAGTTCTGTCTTGGACACTGTGTACTTAGACCAAGAAAG 1148  
QY 621 AsnAspValGluThrThrSerGluThrGlnGluLeuLeuArgThrGluLysValAsnPro 640  
DB 1149 AACGATGTAGAATATTATAGTGAAACCCAGAGCTACTGAGGACAGAAATTTGTTAATCTCT 1208  
QY 641 LeuArgIleThrGlyThrValCysAlaThrLysIleMetLysLeuArgLysIleLeuGlu 660  
DB 1209 CTGAGAATATATGATATGTGTGTGCCACAAAATATGAACTGAGGAAATTTACTTGAA 1268  
QY 661 LysValGluAlaAlaSerGlyPheThrSerGluGluLysAspProGluGluPheLeuAsn 680  
DB 1269 AAGTGGAGGCTGCATCAGATTTTACTCTGAGAGAAAAGATCTCTGAGGAAATTTTGAAT 1328  
QY 681 IleLeuPheHisIleIleLeuArgValGluProLeuLeuLysIleArgSerAlaGlyGln 700  
DB 1329 ATTCGTTTTCATCATATATTTAAGGGTAGAACCTTTGCTAAAAATAAGATCAGCAGCTCA 1388  
QY 701 LysValGlnAspCysThrPheThrGlnIlePheMetGluLysAsnGluLysValGlyVal 720  
DB 1389 AAGGTACAGATTTGTTACTTCTATCAAAATTTTATGCAAAAAAATGAGAAAGTTGGCGTT 1448  
QY 721 ProThrIleGlnGlnLeuLeuGluTrpSerPheIleAsnSerAsnLeuLysPheAlaGlu 740  
DB 1449 CCCACATTCACAGTGTGTAGATGGTCTTTTATCCACAGTAACCTGAAATTTTGAGAG 1508  
QY 741 AlaProSerCysLeuIleIleGlnMetProArgPheGlyAspPheLysLeuPheLys 760  
DB 1509 GCACCATCATGCTGATTATTACAGATGCTCGATTTGGAAGAGACTTTAAACTATTATAA 1568  
QY 761 LysIlePheProSerLeuGluLeuAsnIleThrAspLeuLeuGluAspThrProArgGln 780  
DB 1569 AAAATTTTCTCTCTGGAATTAATAATAACAGATTTTACTTTGAAGACACTCCACAGAC 1628  
QY 781 CysArgIleCysGlyGlyLeuAlaMetThrGluCysArgGluCysThrAspAspProAsp 800  
DB 1629 TCCCGATATGTGGAGGCTTGCAATGTATGATGTAGAGAAATGCTACCAAGATCCGAC 1688  
QY 801 IleSerAlaGlyLysIleLysGlnPheCysLysThrCysAsnThrGlnValHisLeuHis 820  
DB 1689 ATCTCAGCTGGAAATCAAGCAGATTGTGTAAACCTGCAACACTCAAGTCCACCTTCAT 1748



Db 909 TTTGACGCGGTTTCCAAATCAGATTTGAGCGGTGTAACCTCTTTAGCATTTGGAGGCTACTTA 968  
Qy SerGluValValGluGluAsnThrProProLysMetGluLysGluGlyLeuGluIleMet 580  
Db 969 AGTGAAGTAGTAGAAGAAATACTCCACCAAAATGGAAGAGCCCTGGAGATAATG 1028  
Qy 581 lIeGlyLysLysGlylIeGlnGlyHISrTyAsnSerCysTyLeuAspSerThrLeu 600  
Db 1029 ATTGGGAAGAAAGGCAATCCAGGGTCAATTAACAATTCITGTACTTAGACTCAACCTTA 1088  
Qy 601 PheCysLeuPheAlaPheSerSerValLeuLeuAspThrValLeuLeuArgProLysGluLys 620  
Db 1089 TTTCTGCTTATTGCTTTTGTCTGTCTGGACACTGTGTACTTAGACCCAAAGAAAG 1148  
Qy 621 AsnAspValGluTyTySerGluThrGlnGluLeuLeuArgThrGluIleValAsnPro 640  
Db 1149 AACGATGTAGAAATATATATAGTGAACCCAGAGCTACTGAGGACAGAAATTTGTAATCCT 1208  
Qy 641 LeuArgIleTyTyGlyTyValCysAlaThrLysIleMetLysLeuArgLysIleLeuGlu 660  
Db 1209 CTGAGATATATGATATGTGTGTGCCACAAAATTTATGAACTGAGGAAATACTTGAA 1268  
Qy 661 LysValGluAlaAlaSerGlyPheThrSerGluGluLysAspProGluGluPheLeuAsn 680  
Db 1269 AAGGTGGAGGCTGCATCAGGATTTACCTCTGAAGAAAAAGATCTCTGAGGAATTTCTTGAAT 1328  
Qy 681 lIeLeuPheHISrIleLeuArgValGluProLeuLeuLysIleArgSerAlaGlyGln 700  
Db 1329 ATTCTGTTTTCATCATATTTAAAGGGTAGAACCTTTTCTAAAAATAAGATCAGCAGTCAA 1388  
Qy 701 LysValGlnAspCysTyTyPheTyGlnIlePheMetGluLysAsnGluLysValGlyVal 720  
Db 1389 AGGTACAGATGTGTACTTCTATCAATTTTATGGAATAAAATGAGAAAGTTGGCGTT 1448  
Qy 721 ProThrIleGlnGlnLeuLeuGluTrpSerPheIleAsnSerAsnLeuLysPheAlaGlu 740  
Db 1449 CCCACAATTCAGCAGTGTGTAGATGGTCTTTTATCAACAGTAACCTGAAATTTGGAG 1508  
Qy 741 AlaProSerCysLeuIleGlnMetProArgPheGlyLysAspPheLysLeuPheLys 760  
Db 1509 GCACCATCATGTCTGATTTATTCAGATGCTCGATTTGGAAAAAGACTTTAAACTATTATAA 1568  
Qy 761 LysIlePheProSerLeuGluLeuAsnIleThrAspLeuLeuGluAspThrProArgGln 780  
Db 1569 AAAATTTTCTCTCTGGAATTAATAATATACAGATTTTACTTTGAAGACACTCCACACAG 1628  
Qy 781 CysArgIleCysGlyGlyLeuAlaMetTyTyGluCysArgGluCysTyTyAspAspProAsp 800  
Db 1629 TGCCGGATATGTGGAGGGCTTCCAAATGATGATGATGATGATGATGATGATGATGATGAT 1688  
Qy 801 lIeSerAlaGlyLysIleLysGlnPheCysLysThrCysAsnThrGlnValHISrLeuHIS 820  
Db 1689 ATCTCAGCTGGAAAAATCAAGCAGTTTGTAAAAACCTGCAACACTCAAGTCCACCTTCAT 1748  
Qy 821 ProLysArgLeuAsnHISrTyTyAsnProValSerLeuProLysAspLeuProAspTrp 840  
Db 1749 CCGAAGAGGCTGAATCATAAATATATACCCAGTGTCACTTCCCAAGACTTACCCGACTGG 1808  
Qy 841 AspTrpArgHISrGlyCysIleProCysGlnAsnMetGluLeuPheAlaValLeuCysIle 860  
Db 1809 GACTGGAGACAGCGCTGCATCCCTTGCACGAATATGGAGTATTTGCTGTCTCTGCATA 1868  
Qy 861 GluThrSerHISrTyTyValAlaPheValLysTyTyGlyLysAspSerAlaTrpLeuPhe 880  
Db 1869 GAAACAACCCACTATGTGCTTTTGTGAAGTATGGGAAGACGATTCGCTGGCTCTTC 1928  
Qy 881 PheAspSerMetAlaAspArgAspGlyGlyGlnAnGlyPheAsnIleProGlnValThr 900  
Db 1929 TTTGACAGATGGCCGATGGGATGGTGTGTCAGATGGCTTCAACATTCCTCAAGTCACC 1988  
Qy 901 ProCysProGluValGlyGluTyTyLeuLysMetSerLeuGluAspLeuHISrLeuAsp 920

Db 1989 CCATGCCCCAGAGTAGGAGAGTACTTGAAGATGTCTCTGGAAGACCTGCATTCCTTGGAC 2048  
Qy 921 SerArgArgIleGlnGlyCysAlaAtqArgLeuLeuCysAspAlaTyMetCysMetTy 940  
Db 2049 TCCAGGAGATCCAGGCTGTGCAGGAAGACTGCTTTGTGATGATATATGTGATGTAC 2108  
Qy 941 GlnSerProThrMetSerLeuTyTyLys 949  
Db 2109 CAGAGTCCAAACAATGAGTTTGTACAA 2135  
RESULT 9  
ADB48380  
ID ADB48380 standard; cDNA; 2523 BP.  
XX  
AC ADB48380;  
XX  
DT 04-DEC-2003 (first entry)  
XX  
DE Novel human cDNA SEQ ID NO 290.  
XX  
KW ss; cancer; neurodegenerative disease; human.  
XX  
OS Homo sapiens.  
XX  
PN US2003104529-A1.  
XX  
PD 05-JUN-2003.  
XX  
PF 04-JAN-2002; 2002US-00037270.  
XX  
PR 21-JAN-2000; 2000US-00488725.  
PR 25-APR-2000; 2000US-00552317.  
PR 19-JUL-2000; 2000US-00620312.  
XX  
(ZHOU/) ZHOU P.  
(TANG/) TANG Y T.  
(LIUC/) LIU C.  
(ASUN/) ASUNDI V.  
(DRMA/) DRMANAC R T.  
XX  
Zhou P, Tang YT, Liu C, Asundi V, Drmanac RT;  
WPI; 2003-678194/64.  
XX  
New polynucleotide, useful for treating diseases e.g., cancer or  
neurodegenerative diseases.  
XX  
Claim 1; SEQ ID NO 290; 99pp; English.  
XX  
The invention relates to a polynucleotide comprising a sequence given in  
the specification, or its mature protein-coding portion, or its  
complement. The polynucleotide is useful for treating diseases e.g.,  
cancer or neurodegenerative diseases and many others listed in the  
specification. The present sequence represents a novel human cDNA. Note:  
The sequence data for this patent did not form part of the printed  
CC specification and was obtained in electronic format directly from USPRO  
at seqdata.uspto.gov/sequence.html?DocID=20030104529.  
XX  
SQ Sequence 2523 BP; 757 A; 503 C; 557 G; 706 T; 0 U; 0 Other;  
Alignment Scores:  
Pred. No.: 0 Length: 2523  
Score: 641.00 Matches: 668  
Percent Similarity: 99.55% Conservative: 0  
Best Local Similarity: 99.55% Mismatches: 0  
Query Match: 67.54% Indels: 3  
DB: 9 Gaps: 0  
US-09-671-687A-3 (1-949) x ADB48380 (1-2523)  
Qy 282 LeuCysSerPheAlaCysValGluSerThrIleLeuLeuHISrLeuAsnAspIlePro 301  
Db 129 CTTTGTAGTTTTGCGGTGTTGAAGTACAAATTTCTATTGCAATCATGATATCAATCCCA 188



```
Qy 302 GluSerValThrGlnGluArgProProLysLeuAlaPheMetSerArgGlyValGly 321
Db 189 GAGAGTGTGCGCAGGAAAGGAGGCTCCCAAACTTGCCTTATGTCAAGAGGTGTGGG 248
Qy 322 AspLysGlySerSerSerHisAsnLysProLysAlaThrGlySerThrSerAspProGly 341
Db 249 GACAAAGGTTTCATCCAGTCATAATAAACCAAGGCTACAGGATCTACCTCAGACCCTGGA 308
Qy 342 AsnArg---ArgSerGluLeuPheTyrThrLeuAsnGlySerSerValAspSerGlnPro 360
Db 309 ANTAGAAACAGATCTGAATATTATTTATACCTTAAATGGGTCTTCTGTGACTCACAACA 368
Qy 361 GlnSerLysSerLysAsnThrTrpTyrIleAspGluValAlaGluAspProAlaLysSer 380
Db 369 CAATCCAAATCAAAAAATACATGTGTACATTGATGAAGTTGCAGAGAGCCCTGCAAAATCT 428
Qy 381 LeuThrGluIleSerThrAspPheAspArgSerProProLeuGlnProProVal 400
Db 429 CTTACAGAGATATCTACAGACTTTGACCGTTCTTACCACCACTCCAGCCTCCTCTGTG 488
Qy 401 AsnSerLeuThrThrGluAsnArgPheHisSerLeuProPheSerLeuThrLysMetPro 420
Db 489 AACTCAGTGAACCCGAGAACAGATTCACCTCTTTACCATTTCAGTCTCACCAGATGCC 548
Qy 421 AsnThrAsnGlySerIleGlyHisSerProLeuSerLeuSerAlaGlnSerValMetGlu 440
Db 549 AATACCAATGAAGTATTGGCCACAGTCCACTTCTCTGTCCAGCCAGTCTGTATGGNA 608
Qy 441 GluLeuAsnThrAlaProValGlnGluSerProProLeuAlaMetProProGlyAsnSer 460
Db 609 GAGCTAAACACTGCACCCGCTCAAGAGAGTCCACCCCTTGGCCATGCCCTCTGGGAATCA 668
Qy 461 HisGlyLeuGluValGlySerLeuAlaGluValLysGluAsnProProPheTyrGlyVal 480
Db 669 CATGGCTAGAAAGTGGGCTCATTTGGCTGAAGTTAAGAGAACCCCTCTTCTTATGGGTA 728
Qy 481 IleArgTrpIleGlyGlnProProGlyLeuAsnGluValLeuAlaGlyLeuGluLeuGlu 500
Db 729 ATCCGTTGGATCGTCCAGCCAGGACTGAATGAAGTGTCTCGCTGGACTGGAACTGNA 788
Qy 501 AspGluCysAlaGlyCysThrAspGlyThrPheArgGlyThrArgTyrPheThrCysAla 520
Db 789 GATGAGTGTGAGGCTGTACGATGGAACCTTCAGAGGCACCTCGGTATTTACCTGTGCC 848
Qy 521 LeuLysLysAlaLeuPheValLysLeuLysSerCysArgProAspSerArgPheAlaSer 540
Db 849 CTGAAGAAAGGCGCTGTTTGTGAACCTGAAGAGCTGCAGCCCTGACTCTAGGTTTGCATCA 908
Qy 541 LeuGlnProValSerAsnGlnIleGluArgCysAsnSerLeuAlaPheGlyGlyTyrLeu 560
Db 909 TTGCAGCCGGTTTCCATCAGATTGAGCGCTGTAACTCTTTAGCATTTTGGAGGCTACTTA 968
Qy 561 SerGluValValGluGluAsnThrProProLysMetGluLysGluGlyLeuGluIleMet 580
Db 969 AGTGAAGTAGTAGAAGAAATACTCCACCAAAATGGAAAAGAGGCTTGGAGATAATG 1028
Qy 581 IleGlyLysLysLysGlyIleGlnGlyHisTyrAsnSerCysTyrLeuAspSerThrLeu 600
Db 1029 ATTGGGAAGAAAGAGGATCCAGGCTCATTAACAATCTTGTGTACTTAGACTCAACCTTA 1088
Qy 601 PheCysLeuPheAlaPheSerSerValLeuAspThrValLeuLeuArgProLysGluLys 620
Db 1089 TTTCTGCTATTGCTTTTATGTTCTGTTCTGGACACTGTGTACTTAGACCCCAAGAAAG 1148
Qy 621 AsnAspValGluTyrTyrSerGluThrGlnGluLeuLeuArgThrGluIleValAsnPro 640
Db 1149 AACGATGTAGAAATATTATAGTGAAACCCAGAGCTACTGAGACAGAAATTTGTTAATCT 1208
Qy 641 LeuArgIleTyrGlyTyrValCysAlaThrLysIleMetLysLeuArgLysIleLeuGlu 660
Db 1209 CTGGAATATATGATATGTGTGCCACAAAATTTAGAACTCAGGAAAATATCTTGAA 1268
```

```
Qy 661 LysValGluAlaAlaSerGlyPheThrSerGluGluLysAspProGluGluPheLeuAsn 680
Db 1269 AAGTGGAGGCTGCATCAGGATTTACCTCTGAAGAAAAAGATCTCTGAGAAATCTTGAAT 1328
Qy 681 IleLeuPheHisHisIleIleuArgValGluProLeuLeuLysIleArgSerAlaGlyGln 700
Db 1329 ATTCTGTTTCATCATATTTAAGGGTAGAACCTTTGCTAAAAAATAGATCAGCAGGTCAA 1388
Qy 701 LysValGlnAspCysTyrPheTyrGlnIlePheMetGluLysAsnGluLysValGlyVal 720
Db 1389 AAGGTACAGATTTGTTACTTCTATCAAAATTTTATGGAAAAAATGAGAAAGTTGGCTT 1448
Qy 721 ProThrIleGlnGlnLeuLeuGluTrpSerPheIleAsnSerAsnLeuLysPheAlaGlu 740
Db 1449 CCCACAAATTCAGCAGTGTGTAGATGGTCTTTTATCAACAGTAACTCGAAATTTGCAGAG 1508
Qy 741 AlaProSerCysLeuIleIleGlnMetProArgPheGlyLysAspPheLysLeuPheLys 760
Db 1509 GCACCATCATGTCTGATTTATTCAGATGCTCGATTTGGAAAAAGACTTTAAACATATTAAA 1568
Qy 761 LysIlePheProSerLeuGluLeuAsnIleThrAspLeuLeuGluAspThrProArgGln 780
Db 1569 AAAATTTTCTCTCTGGAATTAATAATACAGATTTACTTGAAGACACTCCAGACAG 1628
Qy 781 CysArgIleCysGlyGlyLeuAlaMetTyrGluCysArgGluCysTyrAspAspProAsp 800
Db 1629 TGCCGGATATGTGGAGGCTTGCATGATGATGATGATGATGATGATGATGATGATGATGAT 1688
Qy 801 IleSerAlaGlyLysIleLysGlnPheCysLysThrCysAsnThrGlnValHisLeuHis 820
Db 1689 ATCTCAGCTGGAAAAATCAAGCAGTTTTGTAAACCTGCACACACTCAAGTCCACCTTCAT 1748
Qy 821 ProLysArgLeuAsnHisLysTyrAsnProValSerLeuProLysAspLeuProAspTrp 840
Db 1749 CCGAAGAGGCTGAATCATATAATATACCCAGTGCATCTCCCAAGACTTTACCCGACTGG 1808
Qy 841 AspTrpArgHisGlyCysIleProCysGlnAsnMetGluLeuPheAlaValLeuCysIle 860
Db 1809 GACTGGAGACAGGCTGCATCCCTGCCAGAAATATGGAGTTATTTGCTGTTCTCTGCATA 1868
Qy 861 GluThrSerHisTyrValAlaPheValLysTyrGlyLysAspAspSerAlaTrpLeuPhe 880
Db 1869 GAAACAAGCCACTATGTGCTTTTGTGAAGTATGGAAGAGCAGATTCGCTCGGCTCTTC 1928
Qy 881 PheAspSerMetAlaAspArgAspGlyGlyGlnAsnGlyPheAsnIleProGlnValThr 900
Db 1929 TTTGACAGATGCGCCGATCGGATGGTGCAGATGGCTTCAACATTCCTCAAGTCACC 1988
Qy 901 ProCysProGluValGlyGluTyrLeuLysMetSerLeuGluAspLeuHisSerLeuAsp 920
Db 1989 CCATGCCCAGAAAGTAGGAGAGTACTTGAAGATGCTCTTGGAAAGACTGCAATTCCTTGGAC 2048
Qy 921 SerArgArgIleGlnGlyCysAlaArgLeuLeuCysAspAlaTyrMetCysMetTyr 940
Db 2049 TCCAGAGAAATCCAAAGGCTGTGCAGAAAGCTGCTTGTGATGTCATATATGTGTCATG 2108
Qy 941 GlnSerProThrMetSerLeuTyrLys 949
Db 2109 CAGAGTCCAACAATGAGTTTGTACAAA 2135
RESULT 10
AAS41617
ID AAS41617 standard; cDNA; 4716 BP.
AC AAS41617;
DT 17-DEC-2001 (first entry)
XX
XX cDNA encoding novel human enzyme polypeptide #833.
KW Human; oxidoreductase enzyme; transferase; hydrolase; lyase; isomerase;
KW ligase; hyperproliferative disorder; immunodeficiency disorder;
KW autoimmune disorder; neurological disorder; metabolic disorder;
```



KW inflammatory disorder; cardiovascular disorder; reproductive disorder;  
KW blood-related disorder; infectious disorder; gene therapy; cytostatic;  
KW anti arthritic; nephrotropic; anticoagulant; ss.

OS Homo sapiens.

XX WO200155301-A2.

XX 02-AUG-2001.

XX 17-JAN-2001; 2001WO-US001239.

XX 31-JAN-2000; 2000US-0179065P.

PR 04-FEB-2000; 2000US-0180628P.

PR 24-FEB-2000; 2000US-0184664P.

PR 02-MAR-2000; 2000US-0186350P.

PR 16-MAR-2000; 2000US-0189874P.

PR 17-MAR-2000; 2000US-0190076P.

PR 18-APR-2000; 2000US-0198123P.

PR 19-MAY-2000; 2000US-0205515P.

PR 07-JUN-2000; 2000US-0209467P.

PR 28-JUN-2000; 2000US-0214886P.

PR 30-JUN-2000; 2000US-0215135P.

PR 07-JUL-2000; 2000US-0216647P.

PR 07-JUL-2000; 2000US-0216880P.

PR 11-JUL-2000; 2000US-0217487P.

PR 11-JUL-2000; 2000US-0217496P.

PR 14-JUL-2000; 2000US-0218290P.

PR 26-JUL-2000; 2000US-0220963P.

PR 26-JUL-2000; 2000US-0220964P.

PR 14-AUG-2000; 2000US-0224518P.

PR 14-AUG-2000; 2000US-0224519P.

PR 14-AUG-2000; 2000US-0225213P.

PR 14-AUG-2000; 2000US-0225214P.

PR 14-AUG-2000; 2000US-0225266P.

PR 14-AUG-2000; 2000US-0225267P.

PR 14-AUG-2000; 2000US-0225268P.

PR 14-AUG-2000; 2000US-0225270P.

PR 14-AUG-2000; 2000US-0225447P.

PR 14-AUG-2000; 2000US-0225757P.

PR 14-AUG-2000; 2000US-0225758P.

PR 14-AUG-2000; 2000US-0225759P.

PR 18-AUG-2000; 2000US-0226279P.

PR 22-AUG-2000; 2000US-0226681P.

PR 22-AUG-2000; 2000US-0226688P.

PR 22-AUG-2000; 2000US-0227182P.

PR 23-AUG-2000; 2000US-0227009P.

PR 30-AUG-2000; 2000US-0228924P.

PR 01-SEP-2000; 2000US-0229287P.

PR 01-SEP-2000; 2000US-0229343P.

PR 01-SEP-2000; 2000US-0229344P.

PR 01-SEP-2000; 2000US-0229345P.

PR 05-SEP-2000; 2000US-0229509P.

PR 05-SEP-2000; 2000US-0229513P.

PR 06-SEP-2000; 2000US-0230437P.

PR 06-SEP-2000; 2000US-0230438P.

PR 08-SEP-2000; 2000US-0231242P.

PR 08-SEP-2000; 2000US-0231243P.

PR 08-SEP-2000; 2000US-0231244P.

PR 08-SEP-2000; 2000US-0231413P.

PR 08-SEP-2000; 2000US-0231414P.

PR 08-SEP-2000; 2000US-0232080P.

PR 08-SEP-2000; 2000US-0232081P.

PR 12-SEP-2000; 2000US-0231968P.

PR 14-SEP-2000; 2000US-0232397P.

PR 14-SEP-2000; 2000US-0232398P.

PR 14-SEP-2000; 2000US-0232399P.

PR 14-SEP-2000; 2000US-0232400P.

PR 14-SEP-2000; 2000US-0232401P.

PR 14-SEP-2000; 2000US-0233063P.

PR 14-SEP-2000; 2000US-0233064P.

PR 14-SEP-2000; 2000US-0233065P.

PR 21-SEP-2000; 2000US-0234223P.

PR 21-SEP-2000; 2000US-0234274P.

PR 25-SEP-2000; 2000US-0234997P.

PR 25-SEP-2000; 2000US-0234998P.

PR 26-SEP-2000; 2000US-0235484P.

PR 27-SEP-2000; 2000US-0235834P.

PR 27-SEP-2000; 2000US-0235836P.

PR 29-SEP-2000; 2000US-0236327P.

PR 29-SEP-2000; 2000US-0236367P.

PR 29-SEP-2000; 2000US-0236368P.

PR 29-SEP-2000; 2000US-0236369P.

PR 29-SEP-2000; 2000US-0236370P.

PR 02-OCT-2000; 2000US-0236802P.

PR 02-OCT-2000; 2000US-0237037P.

PR 02-OCT-2000; 2000US-0237038P.

PR 02-OCT-2000; 2000US-0237039P.

PR 13-OCT-2000; 2000US-0237040P.

PR 13-OCT-2000; 2000US-0239935P.

PR 13-OCT-2000; 2000US-0239937P.

PR 20-OCT-2000; 2000US-0240960P.

PR 20-OCT-2000; 2000US-0241221P.

PR 20-OCT-2000; 2000US-0241785P.

PR 20-OCT-2000; 2000US-0241786P.

PR 20-OCT-2000; 2000US-0241787P.

PR 20-OCT-2000; 2000US-0241808P.

PR 20-OCT-2000; 2000US-0241809P.

PR 20-OCT-2000; 2000US-0241826P.

PR 01-NOV-2000; 2000US-0244617P.

PR 08-NOV-2000; 2000US-0246474P.

PR 08-NOV-2000; 2000US-0246475P.

PR 08-NOV-2000; 2000US-0246476P.

PR 08-NOV-2000; 2000US-0246477P.

PR 08-NOV-2000; 2000US-0246478P.

PR 08-NOV-2000; 2000US-0246523P.

PR 08-NOV-2000; 2000US-0246524P.

PR 08-NOV-2000; 2000US-0246525P.

PR 08-NOV-2000; 2000US-0246526P.

PR 08-NOV-2000; 2000US-0246527P.

PR 08-NOV-2000; 2000US-0246528P.

PR 08-NOV-2000; 2000US-0246532P.

PR 08-NOV-2000; 2000US-0246609P.

PR 08-NOV-2000; 2000US-0246610P.

PR 08-NOV-2000; 2000US-0246611P.

PR 08-NOV-2000; 2000US-0246613P.

PR 17-NOV-2000; 2000US-0249207P.

PR 17-NOV-2000; 2000US-0249208P.

PR 17-NOV-2000; 2000US-0249209P.

PR 17-NOV-2000; 2000US-0249210P.

PR 17-NOV-2000; 2000US-0249211P.

PR 17-NOV-2000; 2000US-0249212P.

PR 17-NOV-2000; 2000US-0249213P.

PR 17-NOV-2000; 2000US-0249214P.

PR 17-NOV-2000; 2000US-0249215P.

PR 17-NOV-2000; 2000US-0249216P.

PR 17-NOV-2000; 2000US-0249217P.

PR 17-NOV-2000; 2000US-0249218P.

PR 17-NOV-2000; 2000US-0249244P.

PR 17-NOV-2000; 2000US-0249245P.

PR 17-NOV-2000; 2000US-0249264P.

PR 17-NOV-2000; 2000US-0249265P.

PR 17-NOV-2000; 2000US-0249297P.

PR 17-NOV-2000; 2000US-0249299P.

PR 17-NOV-2000; 2000US-0249300P.

PR 01-DEC-2000; 2000US-0250160P.

PR 01-DEC-2000; 2000US-0250391P.

PR 05-DEC-2000; 2000US-0251030P.

PR 05-DEC-2000; 2000US-0251988P.

PR 05-DEC-2000; 2000US-0256719P.

PR 06-DEC-2000; 2000US-0251479P.

PR 08-DEC-2000; 2000US-0251856P.

PR 08-DEC-2000; 2000US-0251868P.

PR 08-DEC-2000; 2000US-0251869P.

PR 08-DEC-2000; 2000US-0251989P.

PR 08-DEC-2000; 2000US-0251990P.

PR 11-DEC-2000; 2000US-0254097P.  
 PR 05-JAN-2001; 2001US-0259678P.  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 XX Rosen CA, Baraash SC, Ruben SM;  
 XX WPI; 2001-465566/50.  
 DR P-PSDB; AAU23747.  
 XX  
 PT Novel polypeptides and polynucleotides useful for diagnosing, preventing,  
 PT treating neural, immune system, muscular, reproductive, pulmonary,  
 PT cardiovascular, renal, proliferative disorders and cancerous diseases.  
 XX  
 PS Claim 4; SEQ ID NO 843; 1180pp; English.  
 XX  
 CC The present invention relates to the isolation of novel human enzyme  
 CC polypeptides (AAU22315-AAU23814), and the cDNA and genomic sequences  
 CC encoding them. The enzyme polypeptides of the invention may comprise the  
 CC functional classes of oxidoreductases, transferases, hydrolases, lyases,  
 CC isomerases or ligases. The sequences of the invention are useful in the  
 CC diagnosis, treatment, prevention and/or prognosis of a wide range of  
 CC disorders including hyperproliferative disorders (e.g. cancer),  
 CC immunodeficiency disorders (e.g. AIDS) autoimmune disorders (e.g.  
 CC arthritis), neurological disorders (e.g. Alzheimer's disease), metabolic  
 CC disorders (e.g. phenylketonuria), inflammatory disorders (e.g. asthma),  
 CC cardiovascular disorders (e.g. atherosclerosis), blood-related disorders  
 CC (e.g. haemophilia), reproductive disorders (e.g. infertility) and  
 CC infectious disorders (e.g. influenza). The polynucleotides of the  
 CC invention can also be used in gene therapy. AA540785-AA541684 represent  
 CC cDNA sequences encoding for the novel human enzyme polypeptides of the  
 CC invention. Note: The sequence data for this patent did not form part of  
 CC the printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
 XX

SQ Sequence 4716 BP; 1397 A; 848 C; 1036 G; 1431 T; 0 U; 4 Other;

#### Alignment Scores:

Pred. No.:	0	Length:	4716
Score:	641.00	Matches:	668
Percent Similarity:	99.55%	Conservative:	0
Best Local Similarity:	99.55%	Mismatches:	0
Query Match:	67.54%	Indels:	3
DB:	4	Gaps:	0

US-09-671-687A-3 (1-949) x AA541617 (1-4716)

Qy	282	LeuCysSerPheAlaCysValGluSerThrIleLeuLeuHisIleAsnAspIleIlePro	301
Db	212	CTTTGTAGTTTGGTGTGTTGAAAGTACAAATCTATTGACATCAATGATATCATCCCA	271
Qy	302	GluSerValThrGlnUArgProProLysLeuAlaPheMetSerArgGlyValGly	321
Db	272	GAGAGTGTGACGAGGAAAGGAGGCTCCCAAACTTGGCTTTATGTCAGAGGTGTGG	331
Qy	322	AspLysGlySerSerSerHisAsnLysProLysAlaThrGlySerThrSerAspProGly	341
Db	332	GACAAAGGTTTCATCCAGTCATAATAACCAAGGCTACAGGATCTACCTCAGACCCCTGGA	391
Qy	342	AsnArg---ArgSerGluLeuPheThrThrLeuAsnGlySerSerValAspSerGlnPro	360
Db	392	AATGAAACACATCTGAATATTTTATACCTTAATGGTCTTCCTGTTGACTCACACCA	451
Qy	361	GlnSerLysSerLysAsnThrTrpTrileAspGluValAlaGluAspProAlaLysSer	380
Db	452	CAATCCAAATCAAAAAATACATGTTGATGATGAAGTTGCAGAGAGCCCTGCAAAATCT	511
Qy	381	LeuThrGluIleSerThrAspPheAspArgSerSerProProLeuGlnProProVal	400
Db	512	CTTACAGAGATATCTACAGACTTTGACCGTCTTCACCAACCACTCCAGCCTCCTCTGTG	571
Qy	401	AsnSerLeuThrThrGluAsnArgPheHisSerLeuProPheSerLeuThrLysMetPro	420

Db	572	AACTCACTGACCCAGGAGAAAGATTCACCTCTTTACCATTCAGTCTCCACCAAGATGCC	631
Qy	421	AsnThrAsnGlySerIleGlyHisSerProLeuSerLeuSerAlaGlnSerValMetGlu	440
Db	632	AATACCAATGGAGTATTGGCCACAGTCCACTTTCTCTGTGACCCAGTCTGTAATGGAA	691
Qy	441	GluLeuAsnThrAlaProValGlnGluSerProProLeuAlaMetProProGlyAsnSer	460
Db	692	GAGCTAAACACTGACCCGCTCCAGAGAGTCCACCTTGGCCATGCCTCTGGGAAGTCA	751
Qy	461	HisGlyLeuGluValGlySerLeuAlaGluValLysGluAsnProProPheTrpGlyVal	480
Db	752	CATGCTCTAGAAAGTGGCTCATTTGGCTGAAGTTAAGGAGAACCCCTCTTCTATGGGTA	811
Qy	481	IleArgTrpIleGlyGlnProProGlyLeuAsnGluValLeuAlaGlyLeuGluGlu	500
Db	812	ATCCGTTGGATCGGTGAGCCACAGACTGAATGAAGTGTCTGCTGGACTGGAAGTGA	871
Qy	501	AspGluCysAlaGlyCysThrAspGlyThrPheArgGlyThrArgTrpPheThrCysAla	520
Db	872	GATGAGTGTGAGGCTGTACGGATGGAACCTTCAGAGGCACCTCGGTATTTTCACTGTGC	931
Qy	521	LeuLysLysAlaLeuPheValLysLysSerCysArgProAspSerArgPheAlaSer	540
Db	932	CTGAAGAAAGGCGCTCTTTGTGAAACTGAAGAGCTGAGGCTGACTTAGGTTTGCATCA	991
Qy	541	LeuGlnProValSerAsnGlnIleGluArgCysAsnSerLeuAlaPheGlyGlyTrpLeu	560
Db	992	TTGACGCGGTTTCCAACTCAGATTGAGCGCTGTAACCTCTTTAGCAATTTGGAGGCTACTTA	1051
Qy	561	SerGluValValGluGluAsnThrProProLysMetGluLysGluGlyLeuGluMet	580
Db	1052	AGTGAAGTAGTAGAAGAAATATCTCCACCAAAATGGAAAAAGAGGCTTGGAGATAATG	1111
Qy	581	IleGlyLysLysGlyIleGlnGlyHisTrpAsnSerCysTrpLeuAspSerThrLeu	600
Db	1112	ATTGGGAAGAAAGAGCATCCAGGCTCATTAACAATTTCTTGTACTTAGACTCAACCTTA	1171
Qy	601	PheCysLeuPheAlaPheSerSerValLeuAspThrValLeuLeuArgProLysGluLys	620
Db	1172	TTCTGCTTATTGCTTTTAGTTCTGTTCTGGACACTGTGTTACTTAGACCCCAAGAAAG	1231
Qy	621	AsnAspValGluTrpTrpSerGluThrGlnGluLeuLeuArgThrGluIleValAsnPro	640
Db	1232	AACGATGTAGATATATTATAGTGAACCAAGAGCTACTGAGGACAGAAATTTGTAATCT	1291
Qy	641	LeuArgIleTrpGlyTrpValCysAlaThrLysIleMetLysLeuArgLysIleLeuGlu	660
Db	1292	CTGAGAATATATGGATATGTGTGTGCCACAAAAATTTATGAACCTGAGGAAATACTTGA	1351
Qy	661	LysValGluAlaAlaSerGlyPheThrSerGluGluLysAspProGluGluPheLeuAsn	680
Db	1352	AAAGTGGAGGCTGCATCAGGATTTACCTCTGAAGAAAAAGATCTCTGAGGAAATCTTGA	1411
Qy	681	IleLeuPheHisHisIleLeuArgValGluProLeuLeuLysIleArgSerIleGlyGln	700
Db	1412	ATTCTGTTTCATCATATTTTAAGGGTAGAACCTTTGCTAAAAAATTAAGATCAGACGTC	1471
Qy	701	LysValGlnAspCysTrpPheTrpGlnIlePheMetGluLysAsnGluLysValGlyVal	720
Db	1472	AAGGTACAGATTTGTTACTTCTATCAATTTTATGGAAAAAATAGAGAAAGTTGGCGTT	1531
Qy	721	ProThrIleGlnGlnLeuLeuGluTrpSerPheIleAsnSerAsnLeuLysPheAlaGlu	740
Db	1532	CCCAAAATTCAGCAGTGTGTAGATGGTCTTTTATCAACAGTAACTCTGAAAATTTGCAG	1591
Qy	741	AlaProSerCysLeuIleIleGlnMetProArgPheGlyLysAspPheLysLeuPheLys	760
Db	1592	GCACCATCATGCTCATATTATTCAGATGCTCGATTTGGAAAAAGACTTTAAACATTA	1651
Qy	761	LysIlePheProSerLeuGluLeuAsnIleThrAspLeuLeuGluAspThrProArgGln	780
Db	1652	AAATTTTCTCTCTGGAATTAATATACAGATTTTACTTGAAGACACTCCAGACAG	1711



Db 526 AACAGTTATATCTGTGATGTTTTCCAGGAAAAAGAAAGCTTAGGATATATTTTGGTGT 585  
Qy 265 lAspMetAspAsnProIleGlyAsnTrpAspGlyValArgPheAspGlyVal---LeuCyase 284  
Db 586 GGACATGGATTAACCTTATGGCACTGGGATGGAAGATTTGATGGAGTGCAGCTTTGTAG 645  
Qy 284 rPheAlaCysValGluSerThrIleLeuLeuHisIleAsnAspIleIleProGluSerVa 304  
Db 646 TTTTGGCGTGTCTGAAGTACAAATCTTATGACATCAATGATATCATCCAGAGAGTGT 705  
Qy 304 lThrGlnGluArgProProLysLeuAlaPheMetSerArgGlyValGlyAspLysG1 324  
Db 706 GACCAGAAAGGAGGCGCTCCCAACTTGCCTTTATGTCAAGAGGTGTGGGACAAAAG 765  
Qy 324 ySerSerSerHisAsnLysProLysAlaThrGlySerThrSerAspProGlyAsnArg-- 343  
Db 766 TTCAATCAGTCATTAATAACCAAGGCTACAGGATCTACCTCAGACCCTGGANAATAGAAA 825  
Qy 344 -ArgSerGluLeuPheTyThrLeuAsnGlySerSerValAspSerGlnProGlnSerLy 363  
Db 826 CAGATCTGAATATTTTATACCTTAAATGGGTCTTCTGTGACTCACAAACACCAATCCAA 885  
Qy 363 sSerLysAsnThrTrpTyrlleAspGluValAlaGluAspProAlaLysSerLeuThrG1 383  
Db 886 ATCAAAAAATACATGGTACATGTATGAAGTTGCAGAGACCCTGCAAAATCTCTTACAGA 945  
Qy 383 uIleSerThrAspPheAspArgSerSerProLeuGlnProProProValAsnSerLe 403  
Db 946 GATATCTACAGACTTTGACCGTCTTCCACCACCCTCCAGCGCTCTCTGTGAACTCACT 1005  
Qy 403 uThrThrGluAsnArgPheHisSerLeuProPheSerLeuThrLysMetProAsnThrAs 423  
Db 1006 GACCACCGAGAACAGATTCCTCTTTACCATTCAGTCTCACCAGATGCCCAATACCA 1065  
Qy 423 nGlySerIleGlyHisSerProLeuSerLeuSerAlaGlnSerValMetGluGluLeuAs 443  
Db 1066 TGGAAATATTGGCCACAGCTCCACTTCTCTGTGAGCCAGCTGTGTAATGGAAGAGCTAAA 1125  
Qy 443 nThrAlaProValGlnGluSerProProLeuAlaMetProProGlyAsnSerHisGlyLe 463  
Db 1126 CACTGCAACCCCTCAAGAGAGTCAACCTTGGCCATGGCTCTCGGAACTCACATGGTCT 1185  
Qy 463 uGluValGlySerLeuAlaGluValLysGluAsnProProPheTyThrGlyValIleArgTr 483  
Db 1186 AGAAGTGGGCTCATTTGGCTGAAGTTAAGAGAACCTCTCTTTATGGGTAAATCCGTG 1245  
Qy 483 pIleGlyGlnProProGlyLeuAsnGluValLeuAlaGlyLeuGluLeuGluAspGluCy 503  
Db 1246 GATCGGTGAGCCACAGAGCTGAATGAAAGTGTCTGCTGCACTGGAACTGGAAGATGAGTG 1305  
Qy 503 sAlaGlyCysThrAspGlyThrPheArgGlyThrArgTyThrPheThrCysAlaLeuLysLy 523  
Db 1306 TGCAGGTGTACGATGGAACCTTCAGAGGCACTCGGTATTTCACCTGTGCCCTGAAGAA 1365  
Qy 523 sAlaLeuPheValLysLeuLysSerCysArgProAspSerArgPheAlaSerLeuGlnPr 543  
Db 1366 GGCCTGTGTGAACTGAAGAGCTGCGGCTGACTCTAGGTTTGATCATTCGACCC 1425  
Qy 543 oValSerAsnGlnIleGluArgCysAsnSerLeuAlaPheGlyGlyTyThrLeuSerGluVa 563  
Db 1426 GGTTTCCAATCAGATTGAGCGCTGTAACCTTTAGCATTTGGAGGCTACTTAAAGTGAAGT 1485  
Qy 563 lValGluGluAsnThrProProLysMetGluLysGluGlyLeuGluIleMetIleGlyLy 583  
Db 1486 AGTAGAGAAATAATCTCCACCAAAATGGAAGAGGCTTGGAGATAATGATTGGGAA 1545  
Qy 583 sLysLysGlyIleGlnGlyHisTyAsnSerCysTyThrLeuAspSerThrLeuPheCysLe 603  
Db 1546 GAAGAAAGGATCCAGGGTCAATTAACATCTTGTGTACTTAGACTCAACTTATCTGCTT 1605  
Qy 603 uPheAlaPheSerSerValLeuAspThrValLeu-LeuArgProLysGluLysAsnAspV 623  
Db 1606 ATTTGCTTTTAGTTCTGTCTGGACACTGTGTACTTGA-CCCAAGAAAGAACGATG 1664

RESULT 12

ADQ95919

ID ADQ95919 standard; cDNA; 3311 BP.

XX

Qy 623 alGluTyTyxSextGluThrGlnGluLeuLeuArgThrGluIleValAsnProLeuArgI 643  
Db 1665 TAGAATATATAGTGAACCCAGAGCTACTGAGGACAGAAATTTGTTAATCTCTGAGAA 1724  
Qy 643 lTyThrGlyTyxValCysAlaThrLysIleMetLysLeuArgLysIleLeuGluLysValG 663  
Db 1725 TATATGGATATGTGTGTGCCCAAAATTTATGAAATCTGAGGAAATATCTTGAAGAGTGG 1784  
Qy 663 luAlaAspSerClyPheThrSerGluGluLysAspProGluLysPheLeuAsnIleLeuP 683  
Db 1785 AGGCTGCATCAGGATTTACCTCTGAAGAAAAGATCTCTGAGGAAATCTTGAATATTTCTGT 1844  
Qy 683 heHisIleIleLeuArgValGluProLeuLeuLysIleArgSerAlaGlyGlnLysValG 703  
Db 1845 TTCAATCATATTTTAAGGGTAGAACCTTTGCTTAAANAATAGATCAGCAGGTCAAAAGGTAC 1904  
Qy 703 lAspCysTyxPheTyThrGlnIlePheMetGluLysAsnGluLysValGlyValProThrI 723  
Db 1905 AAGATTGTACTTCTATCAAAATTTTATGGAATAAATGAGAAAGTTGGCGTTCCACAA 1964  
Qy 723 leGlnGluLeuGluTrpSerPheIleAsnSerAsnLeuLysPheAlaGluAlaProS 743  
Db 1965 TTCAGCAGTTGTAGAATGGTCTTTTATCAACAGTAACCTGAAATTTGCAGAGGCCACAT 2024  
Qy 743 erCysLeuIleIleGlnMetProArgPheGlyLysAspPheLysLeuPheLysLysIleP 763  
Db 2025 CATGTCATATTCAGATGCTCCTGATTTGGAAAAGACTTTTAAACTATTTTAAAAAATTT 2084  
Qy 763 heProSerLeuGluLeuAsnIleThrAspLeuLeuGluAspThrProArgGlnCysArgI 783  
Db 2085 TTCTTCTCTGGAATTAATAATAACAGATTTACTTTGAAGACACTCCAGACAGTCCGGA 2144  
Qy 783 leCysGlyGlyLeuAlaMetTyThrGluCysArgGluCysTyxAspAspProAspIleSerA 803  
Db 2145 TATGTGGAGGCTTGAATGTATGATGTAGTGTAGAGAAATGCTACAGCATCCGACATCTCAG 2204  
Qy 803 laGlyLysIleLysGlnPheCysLysThrCysAsnThrGlnValHisLeuHisProLysA 823  
Db 2205 CTGAAAATCAAGCAGTTTTGTAAACCTGCACACTCAAGTCCACTTCAATCCGAGA 2264  
Qy 823 rgLeuAsnHisLysTyxAsnProValSerLeuProLysAspLeuProAspTrpAspTrpA 843  
Db 2265 GGCTGAATCAATAATAATACCCAGTGTCACTTCCCAAGACTTACCCGACTGGAGTGA 2324  
Qy 843 rgHisGlyCysIleProCysGlnAsnMetGluLeuPheAlaValLeuCysIleGluThrS 863  
Db 2325 GACACGGCTGCATCCCTTGCAGAGATATGGAGTTATTTGCTGTTCTCTGCATAGAACAA 2384  
Qy 863 erHisTyxValAlaPheValLysTyxGlyLysAspSerAlaTrpLeuPheAspS 883  
Db 2385 GCCACTATGTGCTTTTGTGAAGTATGGAAGGAGCATTTCTGCCCTGGCTCTCTTTTGA 2444  
Qy 883 erMetAlaAspArgAspGlyGlyGlnAsnGlyPheAsnIleProGlnValThrProCysP 903  
Db 2445 GCATGGCCGATCGGATGTTGCTGAGATGGCTTCAACATTTCTCAAGTCAACCCCATGCC 2504  
Qy 903 roGluValGlyLysLeuLysMetSerLeuGluAspLeuHisSerLeuAspSerArgA 923  
Db 2505 CAGAAAGTAGGAGAGTACTTGAAGATGTTCTCTGGAAGACCTGCATCTCTTGGAGTCCAGGA 2564  
Qy 923 rgIleGlnGlyCysAlaArgArgLeuLeuCysAspAlaTyxMetCysMetTyxGlnSerP 943  
Db 2565 GAATCAAGGCTGTGCAGAAAGACTGCTTTGTGATGTCATATATGTGTCATGTACCAAGATC 2624  
Qy 943 roThrMetSerLeuTyxLys 949  
Db 2625 CAACATGAGTTGTACAAA 2644



Db 1382 TGCAGAACCCCTGCAAAATCTCTTACAGAGATATCTACAGACTTTGACCGTTCTTCACC 1441  
Qy 393 oProLeuGlnProProValAsnSerLeuThrThrGluAsnArgPheHisSerLeuPr 413  
Db 1442 ACCACTCCAGCCTCCTCTGTGAACCTACTGACACCGAGAACAGATCCCACTCTTTACC 1501  
Qy 413 oPheSerLeuThrLysMetProAsnThrAsnGlySerIleGlyHisSerProLeuSerIle 433  
Db 1502 ATTCAAGTCTCAACAGATGCCAAATACCAATGGAAGTAGTTGGCCACAGCTCCACTTTCTCT 1561  
Qy 433 uSerAlaGlnSerValMetGluLeuLeuAsnThrAlaProValGlnGluSerProProle 453  
Db 1562 GTCAAGCCAGTCTGTAAATGGAAGACTAAACACTGACCCGCTCCAGAGAGTCCACCTT 1621  
Qy 453 uAlaMetProProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLysG 473  
Db 1622 GGCCATGCTCTCTGGGAACCTACATGGTCTAGAACTGGGCTCATTTGGCTGAAGTTAAGGA 1681  
Qy 473 uAsnProProPheThrGlyValIleArgTrpIleGlyGlnProProGlyLeuAsnGluVa 493  
Db 1682 GAACCCCTCTTTCTATGGGGTAATCCGTTGGATCGGTGAGCCACAGGACTGAATGAAGT 1741  
Qy 493 lLeuAlaGlyLeuGluLeuGluAspGluCysAlaGlyCysThrAspGlyThrPheArgG 513  
Db 1742 GCTGCTGGACTGGAACTGGAAATGAGTGTGAGGCTGTACGATGGAACCTTCAGAGG 1801  
Qy 513 YThrArgTyrPheThrCysAlaLeuLysLysAlaLeuPheValLysLeuLysSerCysAr 533  
Db 1802 CACTCGGTATTTCACTCTGCGCTGAAGAGCGGCTGTTTGGAACTGAAGAGCTGCAG 1861  
Qy 533 gProAspSerArgPheAlaSerLeuGlnProValSerAsnGlnIleGluArgCysAsnSe 553  
Db 1862 GCCTGACTCTAGGTTTGCACTATTGACGCGGTTTCCAATCAGATTGAGCGCTGTAACTC 1921  
Qy 553 rLeuAlaPheGlyGlyTyrLeuSerGluValValGluGluAsnThrProProLysMetG 573  
Db 1922 TTTAGCATTTGGAGCTACTTAAGTGAAGTAGTAGAAGAAATACTCCACCAAAATGGA 1981  
Qy 573 uLysGluGlyLeuGluIleMetIleGlyLysLysLysGlyIleGlnGlyHisTyrAsnSe 593  
Db 1982 AAAAGAGCTTGGAGATAAGTATGGAGAAAGAAAGGATCCAGGCTCATTAACATTC 2041  
Qy 593 rCysTyrLeuAspSerThrLeuPheCysLeuPheAlaPheSerSerValLeuAspThrVa 613  
Db 2042 TTGTTACTTAGACTCAACTATTCTGCTATTCTGTTTCTTTAGTCTGTTCTGACACTGT 2101  
Qy 613 lLeuLeuArgProLysGlyLysAsnAspValGluTyrTyrSerGluThrGlnGluLeu 633  
Db 2102 GTTACTTAGACCCCAAGAAAGAAAGACGATGTAGAAATATTATAGTGAACCCCAAGAGCTACT 2161  
Qy 633 uArgThrGluIleValAsnProLeuArgIleTyrGlyTyrValCysAlaThrLysIleMe 653  
Db 2162 GAGACAGAAATGTTAATCTCTGAGAAATATATGGATATGTTGTCACCAAAATATAT 2221  
Qy 653 tLysLeuArg-LysIleLeuGluLysValGluAlaAlaSerGlyPheThrSerGluGluL 673  
Db 2222 GAAACTGAA-GAAAACTTGAAGAGGTGGAGGCTGCATCAGGATTTACTCTCGAAGAA 2280  
Qy 673 yAspProGluGluPheLeuAsnIleLeuPheHisIleLeuArgValGluProLeuL 693  
Db 2281 AAGATCTCTGAGGAATTTCTGAATATTTCTGTTTCATCATATTTAAGGGTAGAACCTTTGC 2340  
Qy 693 euLysIleArgSerAlaGlyGlnLysValGlnAspCysTyrPheTyrGlnIlePheMetG 713  
Db 2341 TAAAAATGAAGATCAGCAGGTCAAAAGGTACAGATTGTTACTTCTATCAAAATTTTATGG 2400  
Qy 713 luLysAsnGluLysValGlyValProThrIleGlnGlnLeuLeuGluTrpSerPheIleA 733  
Db 2401 AAAAAATGAGAAAGTGGCGTTCACCAATTCACAGTGTGTAGATGGTCTTTATCA 2460  
Qy 733 snSerAsnLeuLysPheAlaGluAlaProSerCysLeuIleIleGlnMetProArgPheG 753  
Db 2461 ACAGTAACCTGAAATTTGAGAGGACCACTCATGTCGATTATTTCAGATGCCTCGATTG 2520

Qy 753 lyLysAspPheLysLeuPheLysLysIlePheProSerLeuGluLeuAsnIleThrAspL 773  
Db 2521 GAAAGAGACTTTAAACTATTTAAAAAAATTTTTCCTCTCTGGAATTAATAATACAGATT 2580  
Qy 773 euLeuGluAspThrProArgGlnCysArgIleCysGlyGlyLeuAlaMetTyrGluCysA 793  
Db 2581 TACTTTGAAGACACTCCAGACAGTCCGAGATATGTGGAGGGCTTGCATATGTATGATGTA 2640  
Qy 793 tGluCysTyrAspAspProAspIleSerAlaGlyLysIleLysGlnPheCysLysThrC 813  
Db 2641 GGAATGCTACGACCATCCGACATCTCAGCTGGAAAAAATCAAGCAGTTTGTAAACCT 2700  
Qy 813 ysAsnThrGlnValHisLeuHisProLysArgLeuAsnHisLysTyrAsnProValSerL 833  
Db 2701 GCAACACTCAAGTCCACCTTCATCCGAAGAGGCTGAATCAATAATAAATCAATCCAGTGTAC 2760  
Qy 833 euProLysAspLeuProAspTrpAspTrpArgHisGlyCysIleProCysGlnAsnMetG 853  
Db 2761 TTCCCAAGAGACTTACCCGACTGGAGCTGGAGACAGGCTGCATCCCTTGCACAGATATGG 2820  
Qy 853 luLeuPheAlaValLeuCysIleGluThrSerHisTyrValAlaPheValLysTyrGlyL 873  
Db 2821 AGTTATTTGCTGTTCTCTGCATAGAAACAAGCCACTATGTTGCTTTGTGAAGTATGGGA 2880  
Qy 873 ysAspAspSerAlaTrpLeuPhePheAspSerMetAlaAspArgAspGlyGlyGlnAsnG 893  
Db 2881 AGGAGATTTCTGCTGGCTCTCTTTTGACAGATCGCCGATCGGATGGTGTGTCAGATG 2940  
Qy 893 lyPheAsnIleProGlnValThrProCysProGluValGlyGluTyrLeuLysMetSerL 913  
Db 2941 GCTTCAACATCTCTCAAGTCACTCCCATGCCAGAAAGTAGGAGAGTACTTTGAAGATGCTCTC 3000  
Qy 913 euGluAspLeuHisSerLeuAspSerArgArgIleGlnGlyCysAlaArgArgLeuLeuC 933  
Db 3001 TGAAGAGCTGCTATTTCTTGGACTCCAGGAGAAATCCAAGGCTGTGCACGAGACTGCTTT 3060  
Qy 933 ysAspAlaTyrMetCysMetTyrGlnSerProThrMetSerLeuTyrLys 949  
Db 3061 GTGATGCATATATGTCATGTACAGAGTCCAAACATGAGTTGTACAAA 3110

## RESULT 13

AAH18478

ID AAH18478 standard; cDNA; 2569 BP.

XX AAH18478;

XX 26-JUN-2001 (first entry)

XX Human cDNA sequence SEQ ID NO:18586.

DE Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.  
KW Homo sapiens.  
OS EP1074617-A2.  
XX 07-FEB-2001.  
XX 28-JUL-2000; 2000EP-00116126.  
XX 29-JUL-1999; 99JP-00248036.  
XX 27-AUG-1999; 99JP-00300253.  
XX 11-JAN-2000; 2000JP-00118776.  
XX 02-MAY-2000; 2000JP-00183767.  
XX 09-JUN-2000; 2000JP-00241899.  
XX (HELI-) HELIX RES INST.

PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;  
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;  
XX WPI; 2001-318749/34.



XX Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs.

XX Claim 8; SEQ ID NO 18586; 2537pp + Sequence Listing; English.

XX The present invention describes primer sets for synthesizing 5602 full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dT primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesizing polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AAH03166 to AAH13628 and AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632 represent oligonucleotides, all of which are used in the exemplification of the present invention

XX SQ Sequence 2569 BP; 788 A; 470 C; 562 G; 749 T; 0 U; 0 Other;

#### Alignment Scores:

Pred. No.: 0 Length: 2569  
Score: 418.00 Matches: 585  
Percent Similarity: 98.15% Conservative: 0  
Best Local Similarity: 98.15% Mismatches: 1  
Query Match: 44.05% Indels: 11  
DB: 4 Gaps: 0

US-09-671-687A-3 (1-949) x AAH18478 (1-2569)

Qy 88 ValValGluLeuAsnGluLysPheThrGluLeuLeuLeuAlaIleThrAsnCysGluGlu 107  
Db 1 GTTGAGAGATAAATGAAAGATTCACAGATTACTTTTGGCAATTACCAATTGTGAGGAG 60  
Qy 108 ArgPheSerLeuPheLysAsnArgAsnArgLeuSerLysGlyLeuGlnIleAspValGly 127  
Db 61 AGGTTACGCCCTGTTTAAACACAGAACACAGACTAAGTAAGGCTCCAAATAGACGTGGGC 120  
Qy 128 CysProValLysValGlnLeuArgSerGlyGluGluLysPheProGlyValValArgPhe 147  
Db 121 TGTCTGTGMAAGTACAGCTGAGATCTGGGGAAGAAAATTTCTCGAGTTGTACGCTTC 180  
Qy 148 ArgGlyProLeuLeuAlaGluArgThrValSerGlyIlePhePheGlyValGluLeuLeu 167  
Db 181 AGAGACCCCTGTTAGCAGCAGAGGACAGTCTCCGGAATATTCTTTGAGTTGAATTCGTCG 240  
Qy 168 GluGluGlyArgGlyGlnGlyPheThrAspGlyValTyrGlnGlyLysGlnLeuPheGln 187  
Db 241 GAAGAAGTGTGGGTCAAGAGTTTCACTGACGGGGGTGTACCAAGGGAAAACAGCTTTTTCAG 300  
Qy 188 CysAspGluAspCysGly---PheValAlaLeuAspLysLeuGluLeuIleGluAspAsp 206  
Db 301 TGTGATGAAGATTGTGGCGTGTGTTGTCATTGGACAACTAGAACTCATAGAGATGAT 360  
Qy 207 AspThrAlaLeuGluSerAspTyrAlaGlyProGlyAspThrMetGlnValGluLeuPro 226  
Db 361 GACACTGCAATTGGAAGTGATTAACGACGCTCTGGGGACACAAATGACGAGTCGAATTCCT 420  
Qy 227 ProLeuGluIleAsnSerArgValSerLeuLysGly-GlyGluThrIleGluSerGlyTh 246

Db 421 CCTTTGGAATAAATCCAGAGTTTCTTTGAA--GGTTGGAGAAACAATAAGAACTCGAAC 479  
Qy 246 rValIlePheCysAspValLeuProGlyLysGluSerLeuGlyTyrPheValGlyValAs 266  
Db AGTTATATTCTGTGATGTTTGGCAGGAAAGAAAGCTTAGGATATTTTGTGTGTGGA 539  
Qy 266 pMetAspAsnProIleGlyAsnTrpAspGlyArgPheAspGlyVal---LeuCysSerPh 285  
Db 540 CATGGATAAACCTATTGGCAACTGGGATGGAAGATTGATGGAGTGAGCTTTGTAGTTT 599  
Qy 285 eAlaCysValGluSerThrIleLeuHisIleAsnAspIleIleProGluSerValTh 305  
Db 600 TGGCGTGTGTGAAGTACAAATCTATTGCACATCAATGATATCATCCAGAGAGTGTGAC 659  
Qy 305 rGlnGluArgArgProProLysLeuAlaPheMetSerArgGlyValGlyAspLysGlySe 325  
Db 660 GAGGAAAGAGAGCCCTCCCAAACTTGCCCTTATGTCAAGAGGTGTTGGGACAAAGTTTC 719  
Qy 325 rSerSerHisAsnLysProLysAlaThrGlySerThrSerAspProGlyAsnArg---Ar 344  
Db 720 ATCCAGTCATATATACCAAGGCTACAGGATCTACTCAGACCTCGGAATAAGAAACAG 779  
Qy 344 gSerGluLeuPheTyrThrLeuAsnGlySerSerValAspSerGlnProGlnSerLysSe 364  
Db 780 ATCTGAATATTATTATACCTTAAATGGGCTCTTCTGTGTACTCACAACCAACATCCAAATC 839  
Qy 364 rLysAsnThrTyrTyrIleAspGluValAlaGluAspProAlaLysSerLeuThrGluI 384  
Db 840 AAAAAATAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 899  
Qy 384 eSerThrAspPheAspArgSerProProLeuGlnProProProValAsnSerLeuTh 404  
Db 900 ATCTACAGACTTTGACCGTCTTTCACACCACTCCAGCCCTCTCTGTGACTCACTGAC 959  
Qy 404 rThrGluAsnArgPheHisSerLeuProPheSerLeuThrLysMetProAsnThrAsnGl 424  
Db 960 CACCGAGAACAGATTCCACTCTTTTACCAATTCAGTCTCACCAGATGCCCAATACCAATGG 1019  
Qy 424 ySerIleGlyHisSerProLeuSerLeuSerAlaGlnSerValMetGluGluLeuAsnTh 444  
Db 1020 AAGTATTGGCCACAGTCCACTTTCTCTGTGAGCCCACTGTGTATGAGAGGTAAACAC 1079  
Qy 444 rAlaProValGlnGluSerProProLeuAlaMetProGlyAsnSerHisGlyLeuGl 464  
Db 1080 TGACCCCGTCCAGAGAGAGTCCACCTTGGCCATGCGCTCTCTGGGAATCATCATGTCTAGA 1139  
Qy 464 uValGlySerLeuAlaGluValLysGluAsnProProPheTyrGlyValIleArgTrpI 484  
Db 1140 AGTGGGCTCATTTGGCTGAAGTTAAGGAGAACCTCTCTTTCTATGGGGTAATCCGTTGAT 1199  
Qy 484 eGlyGlnProProGlyLeuAsnGluValLeuAlaGlyLeuGluLeuGluAspGluCysAl 504  
Db 1200 CGGTGACCCACCAAGGACTGAATGAAGTGTGCTGCTGAGCTGGAACCTGGAAGATGAGTGTGC 1259  
Qy 504 aglyCysThrAspGlyThrPheArgGlyThrArgTyrPheThrCysAlaLeuLysVal 524  
Db 1260 AGGCTGTACGGATGGAAACCTTCAGAGGCATCGGTATTTCACCTGTGCCCTGGAAGAGGC 1319  
Qy 524 aLeuPheValLysLeuLysSerCysArgProAspSerArgPheAlaSerLeuGlnProVa 544  
Db 1320 GCTGTTTGTGAACCTGAAAGAGCTGAGGCTGACTCTAGGTTTGCATCATTTGACGCGGT 1379  
Qy 544 lSerAsnGlnIleGluArgCysAsnSerLeuAlaPheGlyGlyTyrLeuSerGluValVa 564  
Db 1380 TTCCAATCAGATTGAGCGCTGTAACCTTTTAGCATTTGGAGGCTACTTAAGTGAAGTAGT 1439  
Qy 564 lGluGluAsnThrProProLysMetGluLysGlyLeuGluIleMetIleGlyLysLy 584  
Db 1440 AGAAGAAAAATCTCCACCAAAAAATGGAAGAGAGGCTTGAGATATATGATGGGAGAA 1499  
Qy 584 sLysGlyIleGlnGlyHisTyrAsnSerCysTyrLeuAspSerThrLeuPheCysLeuPh 604  
Db 1500 GAAAGCATCCAGGCTCATTAACATTTCTGTACTTAGACTCAACCTTATTCTGCTTATT 1559



Qy 604 eAlaPheSerSerValLeuAspThrValLeuLeuArgProLysGluLysAsnAspValGl 624  
 Db 1560 TGCCTTTAGTTCTGTCTGGACACCTGTCTACTAGACCAAGAAAGAACGATGTAGA 1619  
 Qy 624 uTyrTyrSerGluThrGlnGluLeuLeuArgThrGluLeuValAsnProLeuArgIleTyr 644  
 Db 1620 ATATTATAGTGAACCCAGAGCTACTGAGGACAGAAATTGTAATCCTCTGAGAAATATA 1679  
 Qy 644 rGlyTyrValCysAlaThrLysIleMetLysLeuArgLysIleLeuGluLysValGluAl 664  
 Db 1680 TGGATATCTGTGTGCCACAAAATTTATGAACTGAGGAAATACTTTGAAAGGTGAGGC 1739  
 Qy 664 aAlaSerGlyPheThrSerGluGluLys 673  
 Db 1740 TGCATCAGGATTTACCTCTGAAGAAAAA 1767  
 RESULT 14  
 AAI60196  
 ID AAI60196 standard; cDNA; 4286 BP.  
 XX  
 AC AAI60196;  
 XX  
 XX  
 DT 22-OCT-2001 (first entry)  
 XX  
 XX Human polynucleotide SEQ ID NO 4185.  
 XX Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;  
 KW peripheral nervous system; neuropathy; central nervous system; CNS;  
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;  
 KW leukaemia; ss.  
 XX  
 XX Homo sapiens.  
 OS  
 XX WO200153312-A1.  
 FN  
 XX 26-JUL-2001.  
 PD  
 XX 26-DEC-2000; 2000WO-US034263.  
 PF  
 XX 23-DEC-1999; 99US-00471275.  
 PR 21-JAN-2000; 2000US-00488725.  
 PR 25-APR-2000; 2000US-00552317.  
 PR 20-JUN-2000; 2000US-00598042.  
 PR 19-JUL-2000; 2000US-00620312.  
 PR 03-AUG-2000; 2000US-00653450.  
 PR 14-SEP-2000; 2000US-00662191.  
 PR 19-OCT-2000; 2000US-00693036.  
 PR 29-NOV-2000; 2000US-00727344.  
 XX (HYSE-) HYSEQ INC.  
 PA  
 XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;  
 PI Zhou P, Goodrich R, Dmanac RT;  
 XX  
 DR WPI; 2001-442253/47.  
 DR P-PSDB; AAM41040.  
 XX  
 XX Novel nucleic acids and polypeptides, useful for treating disorders such  
 PT as central nervous system injuries.  
 PT  
 XX Claim 1; SEQ ID NO 4185; 10078pp; English.  
 FS  
 XX The invention relates to human nucleic acids (AAI57798-AAI61369) and the  
 CC encoded polypeptides (AAM38642-AAM42213) with nootropic,  
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful  
 CC in gene therapy. A composition containing a polypeptide or polynucleotide  
 CC of the invention may be used to treat diseases of the peripheral nervous  
 CC system, such as peripheral nervous injuries, peripheral neuropathy and  
 CC localised neuropathies and central nervous system diseases, such as

CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
 CC utilisation of the activities such as: Immune system suppression,  
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
 CC assays for receptor activity, arthritis and inflammation, leukaemia and  
 CC C.N.S disorders. Note: The sequence data for this patent did not form  
 CC part of the printed specification  
 XX

SQ Sequence 4286 BP; 1279 A; 779 C; 916 G; 1312 T; 0 U; 0 Other;

#### Alignment Scores:

Pred. No.: 0 Length: 4286  
 Score: 402.00 Matches: 653  
 Percent Similarity: 97.75% Conservative: 0  
 Best Local Similarity: 97.75% Mismatches: 1  
 Query Match: 42.36% Indels: 15  
 DB: 4 Gaps: 0

US-09-671-687A-3 (1-949) x AAI60196 (1-4286)

Qy 282 LeuCysSerPheAlaCysValGluSerThrIleLeuLeuHisIleAsnAspIlePro 301  
 Db 128 CTTTGTAGTTTTCGCTGTGTGAAAGTACAATTCTATTGCATCATCATCATCCCA 187  
 Qy 302 GluSerValThrGlnGluArgProLysLeuAlaPheMetSerArgGlyValGly 321  
 Db 188 GAGAGTGTGCGCAGGAAGAGGAGGCTCCCAAACTTGCCTTTATGTCAGAGGTGTGG 247  
 Qy 322 AspLysGlySerSerHisAsnLysProLysAlaThrGlySerThrSerAspProGly 341  
 Db 248 GACAAGGTTTCATCCAGTCATAATAAACAAGGCTACAGGATCTACCTCAGACCTGGA 307  
 Qy 342 AsnArg---ArgSerGluLeuPheTyrThrLeuAsnGlySerSerValAspSerGlnPro 360  
 Db 308 AATGAAAAACAGATCTGAATTTATTTATACCTTAAATGGGTCTCTCTGTGACTCACAACCA 367  
 Qy 361 GlnSerLysSerLysAsnThrTyrTyrIleAspGluValAlaGluAspProAlaLysSer 380  
 Db 368 CAATCCAAATCAAAAATACATGATGATGATGATGATGATGATGATGATGATGATGAT 427  
 Qy 381 LeuThrGluIleSerThrAspPheAspArgSerSerProProLeuGlnProProVal 400  
 Db 428 CTTACAGAGATATCTACAGACTTTTCACCGTCTCTTCCACCACCTCCAGCTCTCTGTG 487  
 Qy 401 AsnSerLeuThrThrGluAsnArgPheHisSerLeuProPheSerLeuThrLysMetPro 420  
 Db 488 AACTCACTGACCCAGCAGACAGATTCACACTCTTTTACCATTTCAGTCTCACCAGATGCC 547  
 Qy 421 AsnThrAsnGlySerIleGlyHisSerProLeuSerLeuSerAlaGlnSerValMetGlu 440  
 Db 548 AATACCAATGAAGTATTGGCCACAGTCCACTTTCTGTGACCCAGTCTGTATATGAA 607  
 Qy 441 GluLeuAsnThrAlaProValGlnGluSerProProLeuAlaMetProProGlyAsnSer 460  
 Db 608 GAGCTAAACACTGCACCCGTCACAGAGATCCACCTTGGCCATGCTCTCTGGGAACCTCA 667  
 Qy 461 HisGlyLeuGluValGlySerLeuAlaGluValLysGluAsnProPheTyrGlyVal 480  
 Db 668 CATGCTCTAGAAGTGGGCTCATTTGGCTGAAGTAAAGAGAACCCCTCTCTTCTATGGGTA 727  
 Qy 481 IleArgTTrpIleGlyGlnProGlyLeuAsnGluValLeuAlaGlyLeuGluLeuGlu 500  
 Db 728 ATCCGTTGGATCGGTGAGCCACAGGACTGAATGAAGTGTCTGCTGGACTGGAACCTGAA 787  
 Qy 501 AspGluCysAlaGlyCys--ThrAspGlyThrPhe-Arg-GlyThrArgTyrPheThrCy 519  
 Db 788 GATGAGTGTGAGGCTG--GTACGGATGGAACCTTCAAGAGGCATCTCGGTATTTACCTG 846  
 Qy 519 sAlaLeuLysLysAlaLeuPheValLysSerCysArgProAspSerArgPheAl 539  
 Db 847 TGCCCTGAAGAAGGCGCTGTTTGTGAACCTGAAGAGCTGCAGGCTGACTCTAGTCTTGC 906

```
QY 539 aSerLeuGlnProValSerAsnGlnIleGluArgCysAsnSerLeuAla-Phe-GlyGly 558
Db 907 ATCAATGACGCGGTTTCCAATCAGATTGAGCGCTGTAACCTCTTTAGCCATTGGGAGGC 966
QY 559 -TyrLeuSerGluValGluGluAsnThrPro--ProLysMet-GluLysGluGlyLe 577
Db 967 TTACTTAAGTAGAGTAGGAGAAATACTCCCAACCAAAATGGGAAAGAAAGCGCTT 1026
QY 577 uGluIleMetIle-GlyLysLysLysGlyIleGlnGlyHisTyrAsnSerCysTyrLeuA 597
Db 1027 GGAGATATGATGGGAGAGAGAAGGCAATCAGGCTCATACAAATCTTGTACTTAG 1086
QY 597 spSerThrLeuPheCysLeuPheAlaPheSerSerValLeuAspThrValLeuLeuArgp 617
Db 1087 ACTCAACCTTATCTGCTTATTTGCTTTTGTAGTTCTGTTCTGGACACTGTGTTACTTAGAC 1146
QY 617 roLysGluLysAsnAspValGluTyrTyrSerGluThrGlnGluLeuLeuArgThrGluI 637
Db 1147 CCNAAAGAAAGAACGATGTAGAAATATATAGTGAACCCCAAGAGCTACTGAGCAGAGAAA 1206
QY 637 leValAsnProLeuArgIleTyrGlyTyrValCysAlaThrLysIleMetLysLeuArgL 657
Db 1207 TGTATTAATCCTCTGAGATATATGATATGTTGTGTGTCACAAAATTTATGAACACTGAGGA 1266
QY 657 ysIleLeuGluLysValGluAlaAspSerGlyPheThrSerGluGluLysAspProGluG 677
Db 1267 AAATACTTGAAGAGGTGGAGGCTGCATCAGGATTTACCTCTGNAGAAAAGATCTTGAGG 1326
QY 677 luPheLeuAsnIleLeuPheHisIleLeuArgValGluProLeuLeuLysIleArgS 697
Db 1327 AATCTTGAATATCTGTTTTCATCATATTTAAGGAGTAGAACCTTTGCTAAAAATAAGAT 1386
QY 697 erAlaGlyGlnLysValGlnAspCysTyrPheTyrGlnIlePheMetGluLysAsnGluL 717
Db 1387 CAGCAGGTCAAAGGTACAAGATTTGTTCTTATCAAAATTTTATGGAATAAAATGAGA 1446
QY 717 ysValGlyValProThrIleGlnGlnLeuLeuGluTrpSerPheIleAsnSerAsnLeuL 737
Db 1447 AAGTTGGGTTCCCAACATTCAGCAGTGTGTAGATGCTTTTATCAACAGTAACTCGA 1506
QY 737 ysPheAlaGluAlaProSerCysLeuIleIleGlnMetProArgPheGlyLysAspPheL 757
Db 1507 AATTTGCAGAGGCACCATCATGCTGATTTATTATTCAGATGCTCGATTTGGAAAAGACTTTA 1566
QY 757 ysLeuPheLysIlePheProSerLeuGluLeuAsnIleThrAspLeuLeuGluAspT 777
Db 1567 AACTATTAAAAAATTTTCTCTCTGGAATTAATAATATAACAGATTTTACTTGAAGACA 1626
QY 777 hrProArgGlnCysArgIleCysGlyGlyLeuAlaMetTyrGluCysArgGluCysTyrA 797
Db 1627 CTCCAGACAGTCCGGATATGTGGAGGCTTGCAATGATGATGAGTGTAGAGATGCTACG 1686
QY 797 spAspProAspIleSerAlaGlyLysIleLysGlnPheCysLysThrCysAsnThrGlnV 817
Db 1687 ACGATCCGACATCTCAGCTGGAANAATCAACAGATTTTGTAAACCTCCACACTCAAG 1746
QY 817 alHisLeuHiPProLysArgLeuAsnHisLysTyrAsnProValSerLeuProLysAspL 837
Db 1747 TCCACCTTCATCCGAAGAGGCTGAATATAAATATAAACCAGTGTCACTTCCCAAGACT 1806
QY 837 euProAspTrpAspTrpArgHisGlyCysIleProCysGlnAsnMetGluLeuPheAlav 857
Db 1807 TACCGACTGGACGTGGAGACACGGCTGCATCCCTTGCAGAAATATGGAGTATTATTCGTG 1866
QY 857 alLeuCysIleGluThrSerHisTyrValAlaPheValLysTyrGlyLysAspAspSerA 877
Db 1867 TTCTCTGATAGAAACAGCCACTATGTTGTTGTGAGTATGGGAGGAGCATTCG 1926
QY 877 laTrpLeuPhePheAspSerMetAlaAspArgAspGlyGlnAsnGlyPheAsnIleP 897
Db 1927 CCTGGCTCTTCTTTACAGCAGTGGCCGATCGGATGGTGTGAGATGGCTTCAACATTC 1986
QY 897 roGlnValThrProCysProGluValGlyGluTyrLeuLysMetSerLeuLeuAspLeuH 917
```

```
Db 1987 CTCAAGTCACTCCCATGCCAGAGTAGGAGTACTTGAAGATGTCTCTGGAAGACCTGC 2046
QY 917 isSerLeuAspSerArgArgIleGlnGlyCysAlaArgArgLeuLeuCysAspAla 935
Db 2047 ATTCCTTGGACTCCAGAGAAATCCAAAGCTGTGCACGAAGACTGCTTTTGTGATGCC 2102

RESULT 15
ABL89642/c
ID ABL89642 standard; cDNA; 1151 BP.
XX
AC ABL89642;
XX
DT 24-MAY-2002 (first entry)
XX
DE Human polynucleotide SEQ ID NO 204.
XX
KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
KW anti-allergic; hepatotropic; antidiabetic; anti-inflammatory; anti-ulcer;
KW vulnery; anticonvulsant; antibacterial; antifungal; antiparasitic;
KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
KW neurological disease; infection; human; secreted protein; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200190304-A2.
XX
PD 29-NOV-2001.
XX
PF 18-MAY-2001; 2001WO-US016450.
XX
PR 19-MAY-2000; 2000US-0205515P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Birse CE, Rosen CA;
DR WPI: 2002-122018/16.
DR P-PSDB; ABB89233.
XX
Novel 1405 isolated polypeptides, useful for diagnosis, treatment and
prevention of neural, immune system, muscular, reproductive,
gastrointestinal, pulmonary, cardiovascular, renal and proliferative
disorders.
Claim 4; SEQ ID NO 204; 2081pp + Sequence Listing; English.
The invention relates to novel genes (ABL89449-ABL90853) and proteins
(ABB89040-ABB90444) useful for preventing, treating or ameliorating
medical conditions e.g. by protein or gene therapy. The genes are
isolated from a range of human tissues disclosed in the specification.
The nucleic acids, proteins, antibodies and (ant)agonists are useful in
the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
ovarian cancer and other cancers of the adrenal gland, bone, marrow,
breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune
disorders e.g. Addison's disease, allergies, autoimmune haemolytic
anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
cardiovascular disorders such as myocardial ischaemias; (d) wound healing
; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)
infectious diseases such as viral, bacterial, fungal and parasitic
infections. Note: The sequence data for this patent did not form part of
the printed specification, but was obtained in electronic format directly
from WIPO at ftp.wipo.int/pub/published_pct_sequences
SQ Sequence 1151 BP; 320 A; 248 C; 219 G; 355 T; 0 U; 9 Other;

Alignment Scores:
Pred. No.: 1.29e-275 Length: 1151
Score: 279.00 Matches: 279
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 29.40% Indels: 0
```

```
DB: 6 Gaps: 0
US-09-671-687A-3 (1-949) x ABL89642 (1-1151)
Qy 582 GlyLysLysGlyGlyLeuGlnGlyHisTyrAsnSerCysTyrLeuAspSerThrLeuPhe 601
Db 996 GGGAGAGAGAGAGAGATCCAGGTCATTACAAATCTTGTTACTTAGACTCAACCTTATTC 937
Qy 602 CysLeuPheAlaPheSerValLeuAspThrValLeuLeuArgProLysGlyLysAsn 621
Db 936 TGCATTATTTGCTTTTGTCTGTCGACACTGTGTTACTTAGACCCCAAGAGAAAGAAC 877
Qy 622 AspValGluTyrTyrSerGluThrGlnGluLeuLeuArgThrGluIleValAsnProLeu 641
Db 876 GATGTAGAATATTATAGTGAACCCAGAGCTACTCAGGACAGAAATTTGTTAATCCTCTG 817
Qy 642 ArgIleTyrGlyTyrValCysAlaThrLysIleMetLysLeuArgLysIleLeuGlyLys 661
Db 816 AGAATATATGATATGTGTGTCACCAAAATTTATGAAACTGAGGAAATACTTGAAGAAG 757
Qy 662 ValGluAlaAlaSerGlyPheThrSerGluGluLysAspProGluGluPheLeuAsnIle 681
Db 756 GTGGAGGCTGCATCAGGATTTACCTCTGAAGAGAAAGATCCAGAGAAATTTTGATATT 697
Qy 682 LeuPheHisHisIleLeuArgValGluProLeuLysIleArgSerAlaGlyGlnLys 701
Db 696 CTGTTTCATCATATTTTAAGGTAGAACCTTTGCTAAAAATAAGATCAGCAGGTCAAAAG 637
Qy 702 ValGlnAspCysTyrPheTyrGlnIlePheMetGluLysAsnGluLysValGlyValPro 721
Db 636 GTACAAGATTGTACTTCTATCAAAATTTTATGGAAGAAAAATGAGAAAGTTGGCGTTCCC 577
Qy 722 ThrIleGlnGlnLeuLeuGluTyrPheIleAsnSerAsnLeuLysPheAlaGluAla 741
Db 576 ACAATTCAGCAGTGTGTAGATGGTCTTTTATCAACAGTAACCTGAAATTTGCAGAGCA 517
Qy 742 ProSerCysLeuIleGlnMetProArgPheGlyLysAspPheLysLeuPheLysLys 761
Db 516 CCATCATGCTGATTATTTCAGATGCTCGATTTGGAAAAAGACTTTTAAACTATTTAAAAA 457
Qy 762 IlePheProSerLeuGluLeuAsnIleThrAspLeuLeuGluAspThrProArgGlnCys 781
Db 456 ATTTTTCCTCTCTGGAATTAATAATAACAGATTTTACTTGAAGACACTCCCGACAGTGC 397
Qy 782 ArgIleCysGlyGlyLeuAlaMetTyrGluCysArgGluCysTyrAspAspProAspIle 801
Db 396 CGGATATGTGGAGGCTTGCAATGATGAGTGTAGAGAAATGCTACGAYGATCCGGACATC 337
Qy 802 SerAlaGlyLysIleLysGlnPheCysLysThrCysAsnThrGlnValHisLeuHisPro 821
Db 336 TCAGCTGGAAGAAATCAAGCAGTTTGTAAACCTGCAACACTCAAGTCCACCTTCATCGG 277
Qy 822 LysArgLeuAsnHisLysTyrAsnProValSerLeuProLysAspLeuProAspTrpAsp 841
Db 276 AAGAGGCTGAATCATAAATAAACCAGTGTCCACTTCCCAAGACTTACCCGACTGGGAC 217
Qy 842 TrpArgHisGlyCysIleProCysGlnAsnMetGluLeuPheAlaValLeuCysIle 860
Db 216 TGGAGACACGGCTGCATCCCTTGCAGAAATATGGAGTTATTTGCTGTTCTCTGCATA 160
```

Search completed: April 17, 2005, 07:25:19  
Job time : 1183 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: April 17, 2005, 01:26:12 ; Search time 8968 Seconds

(without alignments)  
5127.565 Million cell updates/sec

Title: US-09-671-687A-3

Perfect score: 949

Sequence: 1 MSSGLWSQEKVTSPLYWEERI.....RLLCDAYCMQYQPTMSLYK 949

Scoring table:

Xgapop 60.0 , Xgapext 60.0  
Ygapop 60.0 , Ygapext 60.0  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 4708233 seqs, 24227607955 residues

Word size: 1

Total number of hits satisfying chosen parameters: 9408497

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame\_plus\_p2n.model -DEV=xlp  
-Q=/cgn2\_1/USPTO\_spool\_p/US09671687/runat\_15042005\_163038\_11950/app\_query.fasta\_1.1095  
-DB=GenEmbl -QFMT=fastap -SUFFIX=oligo.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi -LIST=45  
-DOCALLIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc  
-NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09671687 @CGN\_1\_5932 @runat\_15042005\_163038\_11950 -NCPU=6 -ICPU=3  
-NO MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOPOP=6  
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

GenEmbl.\*

1: gb\_ba.\*  
2: gb\_htg.\*  
3: gb\_in.\*  
4: gb\_on.\*  
5: gb\_ov.\*  
6: gb\_pat.\*  
7: gb\_ph.\*  
8: gb\_pl.\*  
9: gb\_pr.\*  
10: gb\_ro.\*  
11: gb\_sts.\*  
12: gb\_sy.\*  
13: gb\_un.\*  
14: gb\_vi.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	754	79.5	3302	6	CQ834224
2	754	79.5	3540	9	BC012342 Homo sapi
3	754	79.5	5414	9	AB020656 Homo sapi
4	711	74.9	4527	6	BD231207 Human cyt

5	685	72.2	5371	9	HSA250014
6	667	70.3	3302	6	CQ834222
7	641	67.5	2523	6	AR338799
8	640	67.4	3480	6	CQ719792
9	608	64.1	2845	6	BD160617
10	608	64.1	2845	6	AX883937
11	608	64.1	2845	9	AK024348
12	598	63.0	3311	6	CQ834226
13	517	54.5	2341	9	AK000187
14	418	44.0	2569	6	BD160470
15	418	44.0	2569	6	AX883681
16	418	44.0	2569	9	AK024212
17	374	39.4	1954	9	AK056226
18	165	17.4	2116	6	BD135433
19	165	17.4	2116	6	CQ867781
20	165	17.4	2116	6	AR534837
21	165	17.4	2116	6	AX017216
22	140	14.8	141663	2	AC145018
23	140	14.8	155857	2	AC145174
24	140	14.8	163319	9	HSA303140
25	140	14.8	168271	9	AC007728
26	140	14.8	178790	2	AC145238
27	124	13.1	837	6	BD149741
28	124	13.1	837	6	AX869679
29	123	13.0	758	6	BD147076
30	123	13.0	758	6	AX867014
31	116	12.2	3267	10	BC082001
32	116	12.2	4314	10	AK122389
33	116	12.2	4501	10	BC042438
34	109	11.5	241990	2	AC123449
35	109	11.5	251132	2	AC098162
36	108	11.4	251957	2	AC126867
37	107	11.3	2426	10	BC049879
38	102	10.7	309	6	CQ055561
39	102	10.7	309	6	CQ074825
40	102	10.7	309	6	CQ105734
41	102	10.7	309	6	CQ144418
42	102	10.7	309	6	CQ179895
43	102	10.7	309	6	CQ204286
44	102	10.7	309	6	CQ265785
45	102	10.7	309	6	CQ302867

ALIGNMENTS

RESULT 1	CQ834224	3302 bp	DNA	linear	PAT 29-JUL-2004
LOCUS	CQ834224	Sequence 95 from Patent WO2004058805.			
DEFINITION	CQ834224	GI:50833761			
ACCESSION	CQ834224.1	GI:50833761			
VERSION	CQ834224.1	GI:50833761			
KEYWORDS	CQ834224.1	GI:50833761			
SOURCE	CQ834224.1	GI:50833761			
ORGANISM	CQ834224.1	GI:50833761			
REFERENCE	CQ834224.1	GI:50833761			
AUTHORS	CQ834224.1	GI:50833761			
TITLE	CQ834224.1	GI:50833761			
JOURNAL	CQ834224.1	GI:50833761			
FEATURES	CQ834224.1	GI:50833761			
source	CQ834224.1	GI:50833761			
CDS	CQ834224.1	GI:50833761			

GenBank accession numbers: HSA250014, CQ834222, AR338799, CQ719792, BD160617, AX883937, AK024348, CQ834226, AK000187, BD160470, AX883681, AK024212, AK056226, BD135433, CQ867781, AR534837, AX017216, AC145018, AC145174, HSA303140, AC007728, AC145238, BD149741, AX869679, BD147076, AX867014, BC082001, AK122389, BC042438, AC123449, AC098162, AC126867, BC049879, CQ055561, CQ074825, CQ105734, CQ144418, CQ179895, CQ204286, CQ265785, CQ302867.

I T N C E R P S L F K A N R N L S K G L Q I D V G C P K V Q L R S G E E K P G V V R P R G P L L A B T V S G  
I P F V E L L E E R G G O T D G V Y Q V G Q L Q C O D E C D G V F A L D K L E L I D D D T A L E S D Y A G  
P G T M Q V L E T N S R V L S K V G E T I B S G T D V C F D L P G K S G J V P G V M D M P G N I W  
D G R F D Q S G F A C V E T I L L H I N D I P E S V T Q E R R P K L A F M S R G V G D Q G S S H N K P  
K A T G S T D Q P G N R N S E L F T Y L I N G S V D S Q S K S N W K W I D E A V D P A K S L E T I S D F  
D R S S P L Q P P V N S L T T E N R H S L F S L T K M N T W S G I G S P L S I S A Q S V M E E L N T A P  
Q E S P L A M P G N S G H L E V S A U S V K E N P F F Y G I W T Q P G L N E M A G L E L D E S C A  
C G T D T F R G T F Y C A L K K L F V L K S C R D S R F A S L Q V S N Q I E R N S L A T F G Y L S E  
V B E N T P M K E Q L L E I M G K K T Q G H Y N C S I G T D T C L F A S P S V L T V L R P K E  
W D E N Y S E T Q E L L E T I N P L R I Y G V C A T I M K L R L K I L E V A A S G F T S E E K D P E E F  
L N I L F H I L R V E P L L I B S A Q K V O D C Y F Q I P M E K N E K V G P T T Q O L E W S F I N S L I  
K F A E A P S G L I I O M P R F G K D F K L I F S L E L N I T D L L E D T P R Q R C I C G L A M Y C R E  
C Y P D D I S A G K I Q C K T C N T Q V H L P K R L N H N P V S L P O L P D W M D R H G C I P C Q N  
E F A V L C I E T H Y V A F Y G K D D S A W L F Y S D N A D R G G Q N G N I P O V T P C P E V G E Y L K  
M S L E D H S L D S R R I O G C A R R L L C D A Y M F O S Q S W T S L Y K"

## ORIGIN

Alignment Scores:

Pred. No.:	0	Length:	3302
Score:	754.00	Matches:	948
Percent Similarity:	98.54%	Conservative:	0
Best Local Similarity:	98.54%	Mismatches:	1
Query Match:	79.45%	Indels:	14
DB:	6	Gaps:	0

US-09-671-687A-3 (1-949) x C0834224 (1-3302)

Qy	1	MetSerSerGlyLeuTrpSerGlnGluIysValThrSerProTyrTrpGluArgIle	20
Db	243	ATGAGTTTCAGGCTTATGAGGCCAAGAAAAGTCACCTTCCACCTACTGGGAAGAGCGGATT	302
Qy	21	PheTyrLeuLeuLeuGlnGluCysSerValThrAspLysGlnThrGlnLysLeuLeuLys	40
Db	303	TTTTTACTTGGCTTCTTCAAGAAATGCAAGCGTTTACAGACAAACAAACAAAGCTCTCTTAAA	362
Qy	41	ValProLysGlySerIleGlyGlnTyrIleGlnAspArgSerValGlyHisSerArgIle	60
Db	363	GTACCGAAGGGAAGTATAGGACAGTATATTCAAGATCGTTCTGTGGGCGATTCAAGGATT	422
Qy	61	ProSerAlaLysGlyLysGlyAsnGlnIleGlyLeuLysIleLeuGluGlnProHisAla	80
Db	423	CCTTCTGCAAAAGGCCAAGAAAATCAGATTGGAATTAATAATTTCTAGAGCAACCTCATGCA	482
Qy	81	ValLeuPheValAspGlu--AspValValGluIleAsnGluLysPheThrGluLeuLeu	99
Db	483	GTTCCTCTTTGTTGATGAAGGAAGTGTGTAGAGATAAATGAAGAAGTTTACAGAGTTACTTT	542
Qy	100	LeuAlaIleThrAsnCysGluGluArgPheSerLeuPheLysAsnArgAsnArgLeuSer	119
Db	543	TTGCAATTACCAATTGTGTAGAGAGAGTTTCAGCGCTGTTTTAAAAACAGAAACAGACTAAGT	602
Qy	120	LysGlyLeuGlnIleAspValGlyCysProValLysValGlnLeuArgSerGlyGluGlu	139
Db	603	AAAGGCGCTTCAAATAGACGTGGGCTGTCTCTGTGAAGATACAGCTGAGATCTCTGGGGAAGAA	662
Qy	140	LysPheProGlyValValArgPheArgGlyProLeuLeuAlaGluArgThrValSerGly	159
Db	663	AAATTTCTCGAGTTGTACGCTTCAGAGGACCCCTGTTAGCAGAGAGGACAGTCTCCGGA	722
Qy	160	IlePhePheGlyValGluLeuLeuGluGluArgGlyArgGlyGlnGlyPheThrAspGlyVal	179
Db	723	ATATTCTTTGGAGTTGAATTTGTGGAAGAAGTCGTGTGTCAAGGTTTCACTGACGGGGTG	782
Qy	180	TyrGlnGlyLysGlnLeuPheGlnCysAspGluAspCysGly---PheValAlaLeuAsp	198
Db	783	TACCAAGGGGAACACGCTTTTTCAGTGTGATGAAGATTGTGGCGTGTGTTGTTGTCATTGGAC	842
Qy	199	LysLeuGluLeuIleGluAspAspThrAlaLeuGluSerAspTyrAlaGlyProGly	218
Db	843	AAGCTAGNACTCATAGAAGATGATGACACTGCATTTGGAAAGTGATTTACCGAGTCTCTGGG	902
Qy	219	AspThrMetGlnValGluLeuProProLeuGluIleAsnSerArgValSerLeuLysGly	238

Qy	596	uAspSerThrLeuPheCysLeuPheAlaPheSerValLeuAspThrValLeuLeuAr	616
Db	2042	AGACTCAACCTATTCTGCTTATTTCCTTATTGCTTCTGTCGACACTGTGTACTTAG	2101
Qy	616	gProlysGluLysAsnAspValGluTyrTyrSerGluThrGlnGluLeuLeuArgThrGl	636
Db	2102	ACCCAAAGAAAGACGATGTAGATATATTATAGTGAACCCCAAGAGCTACTGAGGACAGA	2161
Qy	636	uileValAsnProLeuArgGileTyrGlyTyrValCysAlaThrLysIleMetLysLeuAr	656
Db	2162	AATTGTTAATCTCTGAGAATATATGGATATGTGTGCCAATAAATTATGAAACTGAG	2221
Qy	656	gLySileLeuGluLysValGluAlaAsnSerGlyPheThrSerGluGluLysAspProGl	676
Db	2222	GAAATACTTGAAAGAGTGGAGCTGCATCAGATTACTCTGAGAAAGATCCTGA	2281
Qy	676	uGluPheLeuAsnIleLeuPheHisIleLeuArgValGluProLeuLeuLysIleAr	696
Db	2282	GGAATTCCTGGAATATCTGTTTCATCATATTTTAAGGGGTAGAACCTTCTGTAATAAAG	2341
Qy	696	gSerAlaGlyGlnLysValGlnAspCysTyrPheTyrGlnIlePheMetGluLysAsnGl	716
Db	2342	ATCAGCAGTCAAAAGGTACAGATTGTACTTCTATCAAAATTTTATGGAATAAATGA	2401
Qy	716	uLysValGlyValProThrIleGlnGlnLeuLeuGluTyrSerPheIleAsnSerAsnLe	736
Db	2402	GAAAGTGTGGCGTCCCAAAATTCAGAGTGTGTAGAAATGCTCTTTATCAACAGTAACCT	2461
Qy	736	uLysPheAlaGluAlaProSerCysLeuIleIleGlnMetProArgPheGlyLysAspPh	756
Db	2462	GAAATTCGAGAGCCACCATCATGCTGATTATTCAGATGCCTCGATTGGAAGACATT	2521
Qy	756	eLysLeuPheLysLysIlePheProSerLeuGluLeuAsnIleThrAspLeuLeuGluAs	776
Db	2522	TAAACTATTTAAAAAATTTTCCCTCTCTCGAATTAATAATAACAGATTACTTCTGAAGA	2581
Qy	776	pThrProArgGlnCysArgIleCysGlyGlyLeuAlaMetTyrGluCysArgGluCysTy	796
Db	2582	CACCTCCAGACAGTCCCGGATATGGAGGGCTTGCATGTATGAGTGTAGAGAAATGCTA	2641
Qy	796	rAspAspProAspIleSerAlaGlyLysIleLysGlnPheCysLysThrCysAsnThrGl	816
Db	2642	CGACGATCCGACATCTCAGCTGGAAATCAAGCAGTTTGTGTAACCTGCAACTCA	2701
Qy	816	nValHisLeuHisProLysArgLeuAsnHisLysTyrAsnProValSerLeuProLysAs	836
Db	2702	AGTCACCTTCATCCGAAGAGCTCAATCAATAATATACCCAGTGTCACTTCCCAAGA	2761
Qy	836	pLeuProAspTrpAspTrpArgHisGlyCysIleProCysGlnAsnMetGluLeuPheAl	856
Db	2762	CTTACCCGACTGGGACTGGAGACACGGCTGCATCCCTTCCCAAGATATGAGTATTTCG	2821
Qy	856	aValLeuCysIleGluThrSerHisTyrValAlaPheValLysTyrGlyLysAspAspSe	876
Db	2822	TGTTCTCTGCATAGAAACAAGCCACTATGCTTTGTGGAAGTATGGGAAGACCATTC	2881
Qy	876	rAlaTrpLeuPhePheAspSerMetAlaAspArgAspGlyGlyGlnAsnGlyPheAsnIle	896
Db	2882	TGCTTGCTCTCTTTTGACAGATGCCGATCGGATGCTGTCAGAAATGGCTTCAACAT	2941
Qy	896	eProGlnValThrProCysProGluValGlyGluTyrLeuLysMetSerLeuGluAspLe	916
Db	2942	TCCTCAAGTACCCCATGCCAAGAGTAGGAGATCTTGAAGATGCTCTCTGGAAGACCT	3001
Qy	916	uHisSerLeuAspSerArgArgIleGlyCysAlaArgLeuLeuCysAspAlaTy	936
Db	3002	GCAITTCCTTGGACTCCAGAGAAATCCAGGCTGTGCAGGAAGCTCTTTGTGATGCATA	3061
Qy	936	rMetCysMetTyrGlnSerProThrMetSerLeuTyrLys	949
Db	3062	TATGTGCATGTACCAAGATCCCAACATGAGTTTGTACAA	3101
<hr/>			
RESULT 2			
BC012342			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REMARK			
COMMENT			
13540			
Location/Qualifiers			
1. .3540			
/organism="Homo sapiens"			
/mol_type="mRNA"			
/db_xref="taxon:9606"			
/clone="MGC:19923 IMAGE:4552767"			
/tissue_type="Uterus, leiomyosarcoma"			
FEATURES			
source			
Clone distribution: MGC clone distribution information can be found			
through the I.M.A.G.E. Consortium/LLNL at: <a href="http://image.llnl.gov">http://image.llnl.gov</a>			
Series: IRAL P.M.A.T. 29 Row: c Column: 5			
This clone was selected for full length sequencing because it			
passed the following selection criteria: matched mRNA gi: 14165257.			
BC012342			
Homo sapiens			
clone MGC:19923 IMAGE:4552767, complete cds.			
BC012342			
MGC.			
BC012342.1			
GI:15214433			
Homo sapiens (human)			
Homo sapiens			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.			
1 (bases 1 to 3540)			
Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,			
Klauser, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,			
Altshul, S.P., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,			
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,			
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, B.,			
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,			
Sheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,			
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,			
Abramson, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J.,			
Mkernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,			
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,			
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,			
Fahey, J., Helton, E., Kettman, M., Madan, A., Young, A.C., Rodighiero, S.,			
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,			
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,			
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,			
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalusz, D.E.,			
Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.			
Generation and initial analysis of more than 15,000 full-length			
human and mouse cDNA sequences			
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)			
22388257			
12477932			
2 (bases 1 to 3540)			
Strausberg, R.			
Direct Submission			
Submitted (15-AUG-2001) National Institutes of Health, Mammalian			
Gene Collection (MGC), Cancer Genomics Office, National Cancer			
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,			
USA			
NIH-MGC Project URL: <a href="http://mgc.nci.nih.gov">http://mgc.nci.nih.gov</a>			
Contact: MGC help desk			
Email: <a href="mailto:cgapsb-remail.nih.gov">cgapsb-remail.nih.gov</a>			
Tissue Procurement: ATCC			
cDNA Library Preparation: Rubin Laboratory			
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)			
DNA Sequencing by: Genome Sequence Centre,			
BC Cancer Agency, Vancouver, BC, Canada			
<a href="mailto:info@bcgsc.bc.ca">info@bcgsc.bc.ca</a>			
Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,			
Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,			
Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo			
Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven			
Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline			
Schein, Duane Smalusz, Michael Smith, Lorraine Spence, Jeff Stott,			
Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,			
George Yang, Scott Zuyderduyn, Marco Marra.			

```

/clone_lib="NIH_MGC_46"
/lab_host="DH10B-R"
/notes="Vector: pOTB7"
1. .3540
/gene="CYLD"
/notes="synonyms: CDMT, CYLD1, EAC, KIAA0849, HSPC057"
/db_xref="LOCUSID:1540"
277. .3138
/db_xref="MIM:605018"
/genes="CYLD"
/codon_start=1
/product="CYLD protein"
/protein_id="AAH12342.1"
/db_xref="GI:15214434"
/db_xref="LOCUSID:1540"
/db_xref="MIM:605018"
/tranlation="MSSSLMSQEKVTSYWEERIFYLLQLQSCVTDKQTKLLKVPKG
SIGQYIDRSVGHRIIPSAKKGQIGUKILEQPHAVLFVDEKQVVEINEKFTLELLA
ITNCEERPSLPKRNRLSKGLQIDVCPVKQLRSGEEKFPVWFRFGLAERTVSG
IFPVEELLEGRGGQFTDGVYQKQLFCQDEDCGVFVALDKLEIETDDDTALESDYAG
PGDTQVLELPLEINRSVSLKVGETIEGTVIFCDVLPKESLGYFVGVDMDNPIGNW
DGRFDGVLCSFACVESITLHINDIIPESVTOERRPKLAFMSRGVGDGKSSHNKP
KATGSDPDGNNRSELFTYLNSSVDSQPSQSKNTWIDEVAEDPAKSLTEISDTP
DRSPFPQPPVNSLITNREHSUFPFSLTKMPNTNGSIHSPLSLSAQSVMEEINTAP
VQESPPLAMPNGSHGLEVSLAEVKENPPFYVIRWIGQPPGLNEVLAGELEDECA
GCTDGTFRGTFTYFCALKKALFKVLCRSPDRFASLOPVSNOIERCNSLAFGGVISE
VVENTPPKMEKGLIEMIGKKGIQGHYNSCYLDSITLFCIFAFSSVLDVLLRPKEK
NDVYSYETORLLEIKINPLRIYGVYCATKIMKRLKILEKVEASGFTSEKOPPEF
LNILFHLIRVEPLKIRSAQKQVQDCFYQIPMEKNKVGVTIQGLEWSFINSNL
KFAAPSCLIIOMFRFGDKFLFKXIFPSLENTIDLLEDTPRCRCICGLLAMEYECRE
CYDDPDISAGIKTCQNTQVHLHPKLNHKNYPVSLPKDLPDMDWRHGCIPQNM
ELFVLCIETSHYVAFVYKGDQAWLFFDSMDARDGGQNGFNIPQVTPCPEVGBYLK
MSLEDLSLSDRRIQGCARRLLCDAYMCWQSPMSLYK"
745. .867
/misc_feature
/gene="CYLD"
/notes="CAP_GLY; Region: CAP-Gly domain. CAP stands for
cytoskeleton-associated proteins"
/db_xref="CDD:pfam01302"
1726. .2025
/misc_feature
/gene="CYLD"
/notes="NIP100; Region: CGS244, NIP100, Dynactin complex
subunit involved in mitotic spindle partitioning in
anaphase B [Cell division and chromosome partitioning]"
/db_xref="CDD:CGS244"

ORIGIN
Alignment Scores:
Pred. NO.: 0 Length: 3540
Score: 754.00 Matches: 948
Percent Similarity: 98.54% Conservative: 0
Best Local Similarity: 98.54% Mismatches: 1
Query Match: 79.45% Indels: 14
DB: 9 Gaps: 0

US-09-671-687A-3 (1-949) x BC012342 (1-3540)
QY 1 MetSerSerGlyLeuThrSerGlnGluLysValThrSerProTyrTrpGluGluArgIle 20
DB 277 ATGAGTTTCAGGCTTATGGAGCAAGAAAAGTCACTTCCACCTACTGGGAAGAGCGGATT 336
QY 21 PheTyrLeuLeuLeuGlnGluCysSerValThrAspLysGlnThrGlnLysLeuLeuLys 40
DB 337 TTTTACTTGCTTCTCAGAGATGACGCTTACAGACAAACAAACAAAGCTCTCTTAA 396
QY 41 ValProLysGlySerIleGlnTyrIleGlnAspArgSerValGlyHisSerArgIle 60
DB 397 GTACCGAAGGAAGTATAGGACAGTATATCAAGATCGTTCTGTGGGCACTTCAAGGATT 456
QY 61 ProSerAlaLysGlyLysLysAsnGlnIleGlyLeuLysIleLeuGluGlnProHisala 80
DB 457 CCTTCGCAAGGCAAGAAAATCAGATTGGATTAAAAATTTCTAGAGCAACCTCATGCA 516
QY 81 ValLeuPheValAspGlu---AspValValGluIleAsnGluLysPheThrGluLeuLeu 99

```

```

Db 517 GTTCTCTTTGTTGATGAAAAGGATGTTTGTAGAGATAAAATGAAAAGTTTACAGAGTTACTTT 576
QY 100 LeuAlaIleThrAsnCysGluGluArgPheSerLeuPheLeuAsnArgAsnArgLeuSer 119
DB 577 TTGGCAATTCACAAATGTGAGGAGAGGTTTACGCTTGTTTTAAACAGAAAACAGACTAAGT 636
QY 120 LysGlyLeuGlnIleAspValGlyCysProValLysValGlnLeuArgSerGlyGluGlu 139
DB 637 AAAGGCTTCCAAATAGACGTGGCTGTCTGTGAAAGTACAGCTGAGATCTGGGGAAGAA 696
QY 140 LysPheProGlyValValArgPheArgGlyProLeuLeuAlaGluArgThrValSerGly 159
DB 697 AAATTTCTGAGAGTTGTACGCTTACAGAGACCCCTGTTAGCAGAGACAGCTCTCCGGA 756
QY 160 IlePhePheGlyValGluLeuLeuGluGlyArgGlyGlnGlyGlnGlyPheThrAspGlyVal 179
DB 757 ATATCTTTTGAGATTGAATTCGTGGAAGAGGCTGCTGTGCAAGGTTTCACTGACCGGGTG 816
QY 180 TyrGlnGlyLysGlnLeuPheGlnCysAspGluAspCysGly---PheValAlaLeuAsp 198
DB 817 TACCAAGGGAACAGCTTTTTCAGTGTGATGAAGATTGTGGCGTGTTTGTTGCATTGGAC 876
QY 199 LysLeuGluLeuIleGluAspAspThrAlaLeuGluSerAspTyrAlaGlyProGly 218
DB 877 AAGCTAGAACTCATAGAAGATGATGACACTGCTTGGAAAGTGAATTACGCAGGCTCTGGG 936
QY 219 AspThrMetGlnValGluLeuProLeuGluIleAsnSerArgValSerLeuLysGly 238
DB 937 GACCAATGTCAGGTCGAATCTTCTCTTTGGAATAAATCTCAGAGTTTCTTTGAA--GGT 995
QY 239 -GlyGluThrIleGluSerGlyThrValIlePheCysAspValLeuProGlyLysGluSe 258
DB 996 TGGAAACAATAGAACTCTGGAACAGTTATATTCTGTGATGTTTTCAGGAAAGAAAG 1055
QY 258 rIeuGlyTyrPheValGlyValAspMetAspAsnProIleGlyAsnTrpAspGlyArgPh 278
DB 1056 CTTAGGATATTTGTTGTTGTTGGACATGGAATACCCCTATTGCAACTGGGATGGAAGATT 1115
QY 278 eAspGlyVal---LeuCysSerPheAlaCysValGluSerThrIleLeuLeuHisIleAs 297
DB 1116 TGATGGAGTCAGGCTTTTGTAGTTTTCGCTGTTGTAAGATGACAAATCTATTGACATCAA 1175
QY 297 nAspIleLeuProGluSerValThrGlnGluArgArgProProLysLeuAlaPheMetSe 317
DB 1176 TGATATCATCCAGAGAGTGTGACCAAGAAAGAGGAGGCTCCCAAACTTGCCTTTATGTC 1235
QY 317 rArgGlyValGlyAspLysGlySerSerHisAsnLysProLysAlaThrGlySerTh 337
DB 1236 AAGAGGTGTTGGGACAAAGGTTTCATCCAGTCATAATAAACCAAGGCTACAGGATCTAC 1295
QY 337 rSerAspProGlyAsnArg---ArgSerGluLeuPheTyrThrLeuAsnGlySerSerVa 356
DB 1296 CTCAGACCTCGAAAATAGAAAACAGATCTGCAATTTATTTATACCTTAAATGGGTCTCTGT 1355
QY 356 lAspSerGlnProGlnSerLysSerLysAsnThrTrpTyrIleAspGluValAlaGluAs 376
DB 1356 TGACTCAACACCAACCAATCCAAATCAAAAATACATGGTACATTGATGAAGTTGCAAGA 1415
QY 376 pProAlaLysSerLeuThrGluIleSerThrAspPheAspArgSerSerProProLeuG 396
DB 1416 CCCTGCAAAATCTCTTACAGAGATATCTACAGACTTTCAGCGCTTCTTACCACCACTCA 1475
QY 396 nProProValAsnSerLeuThrThrGluAsnArgPheHisSerLeuProPheSerIle 416
DB 1476 GCCTCTCTGTGAATCACTGACCAACGAGAACAGATTTCCACTCTTTACCATTCAGTCT 1535
QY 416 uThrLysMetProAsnThrAsnGlySerIleGlyHisSerProLeuSerLeuSerAlaGl 436
DB 1536 CACCAAGATGCCCAATACCAATGGAAGTATTGGCCACAGTCCACTTTCTGTGTGAGCCCA 1595
QY 436 nSerValMetGluGluLeuAsnThrAlaProValGlnGluSerProProLeuAlaMetPr 456

```



Db 1596 GTCTGTAATGGAGAGCTAAACACATGCGACCCCTCCAAAGAGAGTCCACCCCTTGGCCATGCC 1655  
Qy 456 oProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLysGluAsnProPr 476  
Db 1656 TCCTGGGAACCTACATGGCTTAGAGTGGCTCATTTGGCTGAAGTTAAGGAGAACCTCTC 1715  
Qy 476 oPheTyGlyValIleAArgTTPilleGlyGlnProProGlyLeuAsnGluValLeuAlaGl 496  
Db 1716 TTTCTATGGGGTAATCCGTTGGATCGGTGAGCCACCCAGGACTGAATGAAGTGTGCTGCTGG 1775  
Qy 496 yLeuGluLeuGluAspGluCysAlaGlyCysThrAspGlyThrPheArgGlyThrArgTy 516  
Db 1776 ACTGGAACCTGGAAGATGATGTGCGAGGCTGTACGGATGAAACCTTCAGAGGCACCTCGTA 1835  
Qy 516 rPheThrCysAlaLeuLysLysAlaLeuPheValLysLeuLysSerCysArgProAspSe 536  
Db 1836 TTTCCACTGTGCCCTTGAGAGGCGCTGTTGTGAAACTGAAAGCTGCAGGCTGACTC 1895  
Qy 536 rArgPheAlaSerLeuGlnProValSerAsnGlnIleGluAArgCysAsnSerLeuAlaPh 556  
Db 1896 TAGTGTTCATCATTCAGCCGCTTCCAAATCAGATTGAGCGCTGTAACCTCTTAGCAAT 1955  
Qy 556 eGlyGlyTyrLeuSerGluValValGluGluAsnThrProProLysMetGluLysGluGl 576  
Db 1956 TGGAGGCTACTTAACTGAAGTAGTAGAAGAAATACTCCACCAAAATGMAAAAGAGG 2015  
Qy 576 yLeuGluIleMetIleGlyLysLysGlyIleGlnGlyHisTyrAsnSerCysTyrLe 596  
Db 2016 CTTGAGAGATAATGATTTGGGAAGAAGAGCGATCCAGGCTCATTAACATTTCTTTACTT 2075  
Qy 596 uAspSerThrLeuPheCysLeuPheAlaPheSerValLeuAspThrValLeuLeuAr 616  
Db 2076 AGACTCAACCTATTCTGCTTATTTCCTTTAGTCTGTTCTGGACACTGTGTACTTAG 2135  
Qy 616 gProLysGluLysAsnAspValGluTyrTyrSerGluThrGlnGluLeuLeuArgThrGl 636  
Db 2136 ACCCAAGAAAGAACGATGTAGATAATTATAGTGAACCCCAAGAGCTACTGAGGACAGA 2195  
Qy 636 uIleValAsnProLeuArgIleTyrGlyTyrValCysAlaThrLysIleMetLysLeuAr 556  
Db 2196 AATTGTTAATCCTCTGAGAATAATATGGATATGTGTGCGCCACAAATAATTATGAACCTGAG 2255  
Qy 656 gLysIleLeuGluLysValGluAlaAlaSerGlyPheThrSerGluGluLysAspProGl 676  
Db 2256 GAAATACTTGAAAAGGAGGAGCTGCATCAGATTTCCTCTGAAGAAAAGATCCTGA 2315  
Qy 676 uGluPheLeuAsnIleLeuPheHisIleLeuArgValGluProLeuLysIleAr 696  
Db 2316 GGAATTCCTTGAATATCTGTTTCATCATATTTTAAAGGTAGAACCTTTGCTAAATAAG 2375  
Qy 696 gSerAlaGlyGlnLysValGlnAspCysTyrPheTyrGlnIlePheMetGluLysAsnGl 716  
Db 2376 ATCAGCAGGTCAAAGGTACAGATTGTTACTTCTATCAAAATTTTATGMAAAATAATGA 2435  
Qy 716 uLysValGlyValProThrIleGlnGluLeuGluTyrSerPheIleAsnSerAsnLe 736  
Db 2436 GAAAGTTCGCGTTCACAAATTCAGCAGTGTGTAGAAATGCTTTTATCAACAGTAACCT 2495  
Qy 736 uLysPheAlaGluAlaProSerCysLeuIleIleGlnMetProArgPheGlyLysAspPh 756  
Db 2496 GAAATTCAGAGGACCATCATGCTGATTTATTCAGATGCTCGATTTTGGAAAAGACTT 2555  
Qy 756 eLysLeuPheLysIlePheProSerLeuGluLeuAsnIleThrAspLeuLeuGluAs 776  
Db 2556 TAAACTATTTAAAAAATTTTTCCTCTCTCGAAATTAATAATAACAGATTACTTCTTGAAGA 2615  
Qy 776 pThrProArgGlnCysArgIleCysGlyGlyLeuAlaMetTyrGluCysArgGluCysty 796  
Db 2616 CACTCCAGACAGTCCCGGATATGTGGAGGCTTGCATTTGATGATGTAGAGTAATGCTA 2675  
Qy 796 rAspAspProAspIleSerAlaGlyLysIleLysGlnPheCysValLysThrCysAsnThrGl 816  
Db 2676 CGACGATCCGACATCTCAGCTGGAAAATAATCAAGCGATTTTGTAAACCTTCACACTCA 2735

Qy 816 nValHisLeuHisProLysArgLeuAsnHisLysTyrAsnProValSerLeuProLysAs 836  
Db 2736 AGTCCACCTTCATCCGAAGAGGCTGAATATAATAATAAATCCAGTGTCACTTCCCAAGA 2795  
Qy 836 pLeuProAspTTPAspTTPArgHisGlyCysIleProCysGlnAsnMetGluLeuPheAl 856  
Db 2796 CTTTACCAGCTGGAGCTGGAGACACGCTGCATCCCTTGCAGAAATATGAGATTATTTC 2855  
Qy 856 aValLeuCysIleGluThrSerHisTyrValAlaPheValLysTyrGlyLysAspAspSe 876  
Db 2856 TGTTCCTGTCATAGAAACAGCCCACTATGTTGTTTGAAGTATGGGAAGGACGATTC 2915  
Qy 876 rAlaTTPLeuPhePheAspSerMetAlaAspArgAspGlyGlyGlnAsnGlyPheAsnIl 896  
Db 2916 TGCTGGCTCTCTTTGACAGCATGGCCGATGGGATGGTGTCAGAAATGGCTTCAACAT 2975  
Qy 896 eProGlnValThrProCysProGluValGlyGluTyrLeuLysMetSerLeuGluAspLe 916  
Db 2976 TCCTCAAGTCACCCCATGCCAGAGTAGGAGAGTACTTCAAGATGTCTCTGGAAGACCT 3035  
Qy 916 uHisSerLeuAspSerArgArgIleGlnGlyCysAlaArgArgLeuLeuCysAspAlaTy 936  
Db 3036 GCATTCCTTGGACTCCAGGAGAAATCCAAGGCTGTGCACGAAGACTGCTTGTGATGCATA 3095  
Qy 936 rMetCysMetTyrGlnSerProThrMetSerLeuTyrLys 949  
Db 3096 TATGTGATGTACAGATCCAGATCCAAATGAGTTTGTACAA 3135  
RESULT 3  
LOCUS AB020656 5414 bp mRNA linear PRI 17-MAY-2001  
DEFINITION Homo sapiens mRNA for KIAA0849 protein, partial cds.  
ACCESSION AB020656  
VERSION AB020656.2 GI:14133220  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE 1 (sites)  
AUTHORS Nagase,T., Ishikawa,K., Suyama,M., Kikuno,R., Hirose,M., Miyajima,N., Tanaka,A., Kotani,H., Nomura,N. and Ohara,O.  
TITLE Prediction of the coding sequences of unidentified human genes. XII. The complete sequences of 100 new cDNA clones from brain which code for large proteins in vitro  
JOURNAL DNA Res. 5 (6), 355-364 (1998)  
MEDLINE 99156230  
PUBMED 10048485  
REFERENCE 2 (bases 1 to 5414)  
AUTHORS Ohara,O., Suyama,M., Kikuno,R., Nagase,T. and Ishikawa,K.  
TITLE Direct Submission  
JOURNAL Submitted (02-DEC-1998) Osamu Ohara, Kazusa DNA Research Institute, Laboratory of DNA Technology, Yana 1532-3, Kisarazu, Chiba 292-0812, Japan (E-mail:cdnaiffo@kazusa.or.jp, Tel:+81-438-52-3913, Fax:+81-438-52-3914)  
COMMENT On May 17, 2001 this sequence version replaced gi:4240186.  
FEATURES  
Location/Qualifiers  
1..5414  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="fh04363"  
/sex="male"  
/tissue type="brain"  
/clone\_lib="pBluescriptII SK plus"  
/dev\_stage="adult"  
/note="vector:pBluescriptII SK plus. This sequence is replaced that of hK05904 cDNA as a representative cDNA sequence for KIAA0849"  
1..5414  
/gene="KIAA0849"  
<426..3308  
gene  
CDS

```

/gene="KIAA0849"
/codon_start=1
/product="KIAA0849 protein"
/protein_id="BA074872.2"
/db_xref="GI:14133221"
/translation="VFENVIMSSGLWSQEKVTSYWEERIFYLLOQCSVTDKOTOK
LLKYPKSGISQYIQDRSVCHSRIPSAKQKQIGLKILEQPHAVLFVDEKVDVNEK
FTELLAITNCEERFSLFKNRLUSKGLQIDVGCVKVQLRSGBEKFPGVVFRPPLL
AERTVSGIFFVEELLEERGQFTDVYQKQLFQDEDCGVFVALDKLEIEDDDTA
LESYAGFDTMQVELPPELEINSRVLKVGETIESGTIVFCDVLPKESLGVFVGVDM
DNPIGNWDKRGDQVLQCFACVESTILLIINDIIPESVTOERRPKLAFMSRSGDKG
SSSKPKKATGSTDGPNRNESELYTLNGSSVDOSQOSKSNWYIDEVAEDPAKSL
TEISTDRSSPPLQPPVPVNSLTTEHRFSLPFLSLTKMPNTNGSIGHSPLSAQSVM
BELNTPAQVESPPPLPQNGSHGLEVSLAEVKNPPFYGIWGPQPLNEVLGL
ELEDECACTDPTGRTFYFTCAKALKALFKLSKRPDSRFASLPQVSNQIERCNLA
FGYLSEVSEVENTPMEKEGLEIMIGKKKIQGHNSCYLDSTFLCFAPSSVLDVTS
LLRPKENDVYVSETQELLRTIWNPIRIYGVCACTIMKILKILEVEAASGFTSE
EKDPEEFLNIIIFHILRLVEPLLKRSAGQKQDCYFYQIEMEKNEKVGVPITQILLEW
SFINSNLUKFAEAPSCLIITQMPRFKNDPKFKKIFPSLEINITDLELDPQCRICGGL
AMYECRCYDDPDISAGKIQCFKTCNTQVHLHPKRLNKNYVPSLPKDLDPDWDRHG
CIPQNMELFVLGICTETSHYVAFVKYKGDSDSAWIFFDSMADRDGGQFNIPQVTPCP
EVGEYLKMSLEDLHSLDSRRIQGCARRLCLDAYMCWYQSPTMSLYK"

```

## ORIGIN

## Alignment Scores:

```

Pred. No.:      0      Length:      5414
Score:          754.00    Matches:      948
Percent Similarity: 98.54%    Conservative: 0
Best Local Similarity: 98.54%    Mismatches: 1
Query Match:    79.45%      Indels:    14
DB:              9          Gaps:      0

```

US-09-671-687A-3 (1-949) x AB020656 (1-5414)

```

QY      1 MetSerGlyLeuTrpSerGlnGluLysValThrSerProTyrTrpGluAlaArgIle 20
Db      447 ATGAGTTCAGGCTTATGGAGCACAAGAAAAGTCATTCAACCCTACTGGGAAGAGCGGATT 506

QY      21 PheTyrLeuLeuLeuGlnGluCysSerValThrAspLysGlnThrGlnLysLeuLeuLys 40
Db      507 TTTTACTTGCTTCTCAAGATGACGCTTTACAGACAAACAAACAAAGCTCTCTTAAA 566

QY      41 ValProLysGlySerIleGlyGlnTyrIleGlnAspArgSerValGlyHisSerArgIle 60
Db      567 GTACCGAAGGAGGATATAGGACAGATATATTCAAGATCGTTCTGTGGGGCATTCAGAGATT 626

QY      61 ProSerAlaLysGlyLysLysAsnGlnIleGlyLeuLysIleLeuGluGlnProHisAla 80
Db      627 CCTTCTGCAAAAGGCAAGAAAATCAGATTGGATTAAAAATTTCTAGAGCAACCTCATGCA 686

QY      81 ValLeuPheValAspGlu--AspValValGluIleAsnGluLysPheThrGluLeuLeu 99
Db      687 GTTCTCTTTGTGTGAATAAGGATGTTGTAGAGATAAATGAAAAGTTTACAGAGATTACTT 746

QY      100 LeuAlaIleThrAsnCysGluGluArgPheSerLeuPheLysAsnArgAsnArgLeuSer 119
Db      747 TTGCAATTACCAATTGTAGAGAGGTTACGCTGTTTAAAAACAGAAACAGACTAAGT 806

QY      120 LysGlyLeuGlnIleAspValGlyCysProValLysValGlnLeuArgSerGlyGluGlu 139
Db      807 AAAGGCTTCAAAATAGACGTGGGTGTCCTGTGNAAGTACAGCTGAGATCTGGGAAGNA 866

QY      140 LysPheProGlyValValArgPheArgGlyProLeuLeuAlaGluArgThrValSerGly 159
Db      867 AAATTTCTGTGAGTTGTACGCTTACAGAGGACCCCTGTTTAGCAGAGAGACAGCTCTCCGA 926

QY      160 IlePhePheGlyValGluLeuLeuGluGluArgGlyArgGlyGlnGlyPheThrAspGlyVal 179
Db      927 ATATCTTTGAGTGAATTCGTGAAGAAGGTGCTGTGTCAAGGTTTCACTGACCGGGTG 986

QY      180 TyrGlnGlyLysGlnLeuPheGlnCysAspGluAspCysGly--PheValAlaLeuAsp 198
Db      987 TACCAAGGAAACAGCTTTTTCAGTGTGATGAGAGATTGTGGCGTGTTTGTGTCATTGGAC 1046

```

```

QY      199 LysLeuGluLeuIleGluAspAspThrAlaLeuGluSerAspTyrAlaGlyProGly 218
Db      1047 AAGCTAGAACTCATAGAGATGATGACACTGCTATTGGAAAGTGATTACGCAGGCTCTGGG 1106

QY      219 AspThrMetGlnValGluLeuProLeuGluIleAsnSerArgValSerLeuLysGly 238
Db      1107 GACACAATGCAAGTCCGAACTTCTCTCTTTGGAAAATAAACTCCAGAGTTCTTTTGAA -GGT 1165

QY      239 -GlyGluThrIleGluSerGlyThrValIlePheCysAspValLeuProGlyLysGluSe 258
Db      1166 TGGAGAAACAATAGAACTTGGAAACAGTTATATCTGTGATGTTTGGCCAGGAAAGAAAG 1225

QY      258 rLeuGlyTyrPheValGlyValAspMetAspAsnProIleGlyAsnTrpAspGlyArgPh 278
Db      1236 CTTAGGATATTTTGTGTGTGGACATGATAACCTATTGGCAACTGGGATGGAAGATT 1285

QY      278 eAspGlyVal---LeuCysSerPheAlaCysValGluSerThrIleLeuLeuHisIleAs 297
Db      1286 TGATGGAGTGCAGCTTTGTAGTTTTCGCTGTGTTGAAAGTACAAATTTATTGCACATCAA 1345

QY      297 nAspIleIleProGluSerValThrGlnGluArgArgProProLysLeuAlaPheMetSe 317
Db      1346 TGATATCATCCAGAGAGTGTGACGAGGAAAGGAGGCTCCCAAACTTGCCCTTTATGTC 1405

QY      317 rArgGlyValGlyAspLysGlySerSerSerHisAsnLysProLysAlaThrGlySerTh 337
Db      1406 AAGAGTGTGGGGACAAAGGTTTCATCCAGTCATATAAACCAAGGCTACAGGATCTAC 1465

QY      337 rSerAspProGlyAsnArg---ArgSerGluLeuPheTyrThrLeuAsnGlySerSerVa 356
Db      1466 CTCGACCTCGAAATAGAAACAGATCTGAATTAATTTTATACCTTAAATGGGTCTCTGT 1525

QY      356 lAspSerGlnProGlnSerLysSerLysAsnThrTrpTyrIleAspGluValAlaGluAs 376
Db      1526 TGACTCAACACACCAATCCAAATCAAAAATAACATGGTACATGTAGTGAAGTTGCAAGA 1585

QY      376 pProAlaLysSerLeuThrGluIleSerThrAspPheAspArgSerSerProProLeuGlu 396
Db      1586 CCTCGCAAAATCTCTTACAGAGATATCTACAGACTTTGACCGTCTTACACCACTCCA 1645

QY      396 nProProValAsnSerLeuThrThrGluAsnArgPheHisSerLeuProPheSerIle 416
Db      1646 GCCTCTCTCTGTGAACCTCACTGACCACCGAGAACAGATTCCACTCTTTACCATTCAGTCT 1705

QY      416 uThrLysMetProAsnThrAsnGlySerIleGlyHisSerProLeuSerLeuSerAlaGl 436
Db      1706 CACCAAGATGCCCAATACCAATGGAAGTATTGGCCACAGTCCACTTTCTCTGTACGCCCA 1765

QY      436 nSerValMetGluGluLeuAsnThrAlaProValGlnGluSerProProLeuAlaMetPr 456
Db      1766 GTCTGTAATGGAAGAGCTAAACACTGCACCCGCTCCCAAGAGAGTCCACCCCTGGCCCATGCC 1825

QY      456 oProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLysGluAsnProPr 476
Db      1826 TCCTGGGAACCTCACATGTTCTAGAAAGTGGGCTCATTTGGCTGAAGTTAAGGAGAACCTCC 1885

QY      476 oPheTyrGlyValIleArgTTrpIleGlyGlnProProGlyLeuAsnGluValLeuAlaGl 496
Db      1886 TTTCTATGGGTAAATCCGTGGATCGGTGAGCCACCGAGACTGAATGAAGTCTCGCTGG 1945

QY      496 yLeuGluLeuGluAspGluCysAlaGlyCysThrAspGlyThrPheArgGlyThrArgTy 516
Db      1946 ACTGNACTGGAAGATGAGTGTGAGGCTGTACGGATGGAACCTTTACAGAGGACTCGGTA 2005

QY      516 rPheThrCysAlaLeuLysLysAlaLeuPheValLysLysSerCysArgProAspSe 536
Db      2066 TTTTACCTGTGCCCTGAAAGAGGCGCTGTTTGTGAAACTGAAGAGCTGCAGGCGCTGACTC 2065

QY      536 rArgPheAlaSerLeuGlnProValSerAsnGlnIleGluArgCysAsnSerLeuAlaPh 556
Db      2066 TAGTTTGATCATCTTCAGCCCGGTTTCCATCAGATTGAGCGCTGTAACTCTTTAGCATT 2125

```

556 eGlyGlyTyrLeuSerGluValValGluGluAsnThrProProLysMetGluLysGluGI 576  
Db TGGAGGCTACTTAAAGTGAAGTAGTAGAAGAAAATACCTCACCAAAAATGGAAGAGG 2185  
576 YLeuGluIleMetIleGlyLysLysGlyIleGlnGlyHisTyrAsnSerCysTyrIle 596  
Db CTTGGAGATAATGATTGGGAAGAGAAAGGCAATCCAGGTCATTACAATTTCTGTACTT 2245  
596 uAspSerThrLeuPheCysLeuPheAlaPheSerValLeuAspThrValLeuLeuAr 616  
Db AGACTCAACCTTATCTGCTTATTTGCTTTAGTTCTGTCGACACTGTGTACTTAG 2305  
616 gProLysGluLysAsnAspValGluTyrTyrSerGluThrGlnGluLeuLeuArgThrGI 636  
Db ACCCAAGAAAGAACGATGATAGATATATAGTGAACCCCAAGAGCTACTGAGGACAGA 2365  
636 uIleValAsnProLeuArgIleTyrGlyTyrValCysAlaThrLysIleMetLysLeuAr 656  
Db AATTGTTAATCTCTGAGATATATGATATGTTGTGTGCGCACAAAATATGAAACTGAG 2425  
656 gLysIleLeuLysValGluAlaAlaSerGlyPheThrSerGluGluLysAspProGI 676  
Db GAAATACCTTGAAAGGGTGGAGGCTGCATCAGATTTACCTCTGAAGAAAAGATCCTGA 2485  
676 uGluPheLeuAsnIleLeuPheHisIleLeuArgValGluProLeuLeuLysIleAr 696  
Db GGAATCTTGATATCTGTTTCATCATATTTTAAAGGTAGAACCTTTGCTTAAATAAG 2545  
696 gSerAlaGlyGlnLysValGlnAspCysTyrPheTyrGlnIlePheMetGluLysAsnGI 716  
Db ATCAGAGCTCAAAAGGTACAAGATTGTTACTTCTATCAATTTTATGGAAGAAAATGA 2605  
716 uLysValGlyValProThrIleGlnLeuLeuGluThrSerPheIleAsnSerAsnIle 736  
Db GAAATGTCGGCTTCCCAATTCACAGTGTGTAGAAATGCTTTTATCAACAGTAACCT 2665  
736 uLysPheAlaGluAlaProSerCysLeuIleIleGlnMetProArgPheGlyLysAspPh 756  
Db GAAATTCGAGAGGACCATCATGCTGATTTATTCAGATGCTCGATTTGGAAGACATT 2725  
756 eLysLeuPheLysIlePheProSerLeuGluLeuAsnIleThrAspLeuLeuGluAs 776  
Db TAACTATTAAATAAATTTTCTCTCTCGAATTAATATAACAGATTTACTTGAAGA 2785  
776 pThrProArgGlnCysArgIleCysGlyGlyLeuAlaMetTyrGluCysArgGluCysTy 796  
Db CACTCCCAGACAGTCCGGATATGTGGAGGCTTGCAATGTATGATGATGATGATGCTA 2845  
796 rAspAspProAspIleSerAlaGlyValIleLysGlnPheCysLysThrCysAsnThrGI 816  
Db CGAGCATCCGACATCTCAGCTGGAAAAATCAAGCAGTTTGTAAACCTGCAACACTCA 2905  
816 nValHisLeuHisProLysArgLeuAsnHisLysTyrAsnProValSerLeuProLysAs 836  
Db AGTCACCTTCATCCGAGAGGCTGATCATATAATAATACCCAGCTGCTCTCCCAAGA 2965  
836 pLeuProAspTrpAspTrpArgHisGlyCysIleProCysGlnAsnMetGluLeuPheAl 856  
Db CTTACCCGACTGGGACTGGAGACCGGCTGCATCCCTTGCCAGATATGAGTATTATTC 3025  
856 aValLeuCysIleGluThrSerHisTyrValAlaPheValLysTyrGlyLysAspAspSe 876  
Db TGTTCCTCTGCAATAGAAACCAAGCCTATGCTTTGTTGAAGTATGGGAAGGACGATT 3085  
876 rAlaTrpLeuPhePheAspSerMetAlaAspArgAspGlyGlyGlnAsnGlyPheAsnI 896  
Db TGCCTGGCTCTCTTTTGACAGCATGGCCGATCGGATGGTGTGATGATGGCTTCAACAT 3145  
896 eProGlnValThrProCysProGluValGlyGluTyrLeuLysMetSerLeuGluAspLe 916  
Db TCCTCAAGTCAACCCATGCCAGAGATAGAGAGTACTTGAAGATGCTCTGGAAGACCT 3205  
916 uHisSerLeuAspSerArgArgIleGlnGlyCysAlaArgArgLeuLeuCysAspAlaTy 936

Db 3206 GCATTCCTTGGACTCAGAGAAATCAAGCTGTGCACGAAGACTCTTTGTGATGATA 3265  
Qy 936 rMetCysMetTyrGlnSerProThrMetSerLeuTyrLys 949  
Db 3266 TATGTGCATGTACCAGAGTCCCAACATGATGATTTGTACAAA 3305  
RESULT 4  
LOCUS BD231207 4527 bp DNA linear PAT 17-JUL-2003  
DEFINITION Human cytoskeleton associated proteins.  
ACCESSION BD231207  
VERSION BD231207.1 GI:33040977  
KEYWORDS JP 2002526076-A/9.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 4527)  
Lal,P., Tang,T.Y., Yue,H., Hillman,J.L., Bandman,O., Corley,N.C.,  
Guegler,K.J., Patterson,C., Azimzai,Y. and Baughn,M.R.  
Human cytoskeleton associated proteins  
Patient: JP 2002526076-A 9 20-AUG-2002;  
INCYTE PHARMACEUTICALS INC  
OS Homo sapiens (human)  
PN JP 2002526076-A/9  
PD 20-AUG-2002  
PF 17-SEP-1999 JP 2000574254  
PR 18-SEP-1998 US 60/172226, 27-APR-1999 US 60/131321 PI  
PREETI LAL,TOM Y TANG, HENRY YUE, JENNIFER L HILLMAN,OLGA PI  
BANDMAN,  
PI NEIL C CORLEY,KARL J GUEGLER,CHANDRA PATTERSON,YALDA AZIMZAI,  
PI MARIAH R BAUGHN  
PC C12N15/09,A61K38/00,A61K45/00,A61P1/16,A61P3/00,A61P3/06 PC  
,A61P3/10,A61P5/00,  
PC A61P5/38,A61P7/00,A61P7/06,A61P9/00,A61P9/10,A61P11/06 PC  
,A61P13/00,A61P13/12,  
PC A61P17/06,A61P19/02,A61P19/10,A61P21/04,A61P25/00,A61P25/00,  
PC A61P25/04,  
PC A61P25/14,A61P25/16,A61P25/28,A61P29/00,A61P31/04,A61P31/12,  
PC A61P31/18,  
PC A61P33/00,A61P35/00,A61P35/02,A61P37/00,A61P43/00,C07K14/47,  
PC C07K16/18,  
PC C12N1/15,C12N1/19,C12N1/21,C12N5/10,C12P21/02,C12P21/08 PC  
,C12Q1/68,G01N33/15,  
PC G01N33/50,G01N33/53,G01N33/53,G01N33/56,C12N15/00,C12N5/00,  
PC A61K37/02  
CC Inocyte ID No: 2363327  
FH Key Location/Qualifiers  
FT source 1..4527 /Organism='Homo sapiens (human)'.  
FT Location/Qualifiers  
1..4527  
/organism='Homo sapiens'  
/mol\_type='genomic DNA'  
/db\_xref='taxon:9606'

FEATURES  
source

## ORIGIN

## Alignment Scores:

Alignment Scores:  
Pred. No.: 0 Length: 4527  
Score: 711.00 Matches: 944  
Percent Similarity: 98.44% Conservative: 0  
Best Local Similarity: 98.44% Mismatches: 2  
Query Match: 74.92% Indels: 15  
DB: 6 Gaps: 0

US-09-671-687A-3 (1-949) x BD231207 (1-4527)

Qy 4 GlyLeuTrpSerGlnGluLysValThrSerProTyrTrpGluGluArgIlePheTyrLeu 23  
Db 270 GGCATTATGAGCCCAAGAAAAGTCACTTCACCTCTACTGGAGAGCGGATTTTACTTG 329  
Qy 24 LeuLeuGlnGluCysSerValThrAspLysGlnThrGlnLysLeuLysValProLys 43

Db 330 CTTCTTCAAGAAATGAGCGTTTACAGACAAACAAACAAAGCTCTTAAAGTACCGAA- 388  
Qy 44 GlySerIleGlyGlnTyrIleGlnAspArgSerValGlyHisSerArgIleProSerAla 63  
Db 389 GGAAGTATAGACAGTATATTCAAGATCGTTCTGTGGGGCATTCGAAGATTCCTTCTGCA 448  
Qy 64 LysGlyLysLysAsnGlnIleGlyLeuLysLysLeuGluGlnProHisAlaValLeuPhe 83  
Db 449 AAAGGCAAGAAAAATCAGATTGGATTAAAAATCTAGAGCAACCTCATGTCAGTTCTCTTT 508  
Qy 84 ValAspGlu---AspValValGluIleAsnGluLysPheThrGluLeuLeuLeuAlaIle 102  
Db 509 GTTGATGAAAAGGATGTTGTAGAGATAAATGAAAAGTTTCACAGAGTTACTTTTGGCAATT 568  
Qy 103 ThrAsnCysGluGluArgPheSerLeuPheLysAsnArgAsnArgLeuSerLysGlyLeu 122  
Db 569 ACCAATTGTGAGGAGAGGTTCCAGCTGTTTAAAAACAGAACAGACTAAGTAAAGGCCTC 628  
Qy 123 GlnIleAspValGlyCysProValLysValGlnLeuArgSerGlyGluGluLysPhePro 142  
Db 629 CAAATAGACGTGGGCTGTCTGTGAAAGTACAGCTGAGATCTGGGGAAGAAAAATTTCT 688  
Qy 143 GlyValValArgPheArgGlyProLeuLeuAlaGluArgThrValSerGlyIlePhePhe 162  
Db 689 GGAGTTGTACCGCTTCAGAGSACCCCTGTTAGCAGAGAGCAGCTCTCCGGAATATTCTTT 748  
Qy 163 GlyValGluLeuLeuGluGluGlyArgGlyGlnGlyPheThrAspGlyValTyrGlnGly 182  
Db 749 GGAGTTGAATTCGTGGAAGAAGTCTGTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 808  
Qy 183 LysGlnLeuPheGlnCysAspGluAspCysGly---PheValAlaLeuAspLysLeuGlu 201  
Db 809 AAACAGCTTTTTCAGTGTGATGAAGATTGTGCGGTTTGTGTCATTGGACAAGCTAGAA 868  
Qy 202 LeuIleGluAspAspAspThrAlaLeuGluSerAspTyrAlaGlyProGlyAspThrMet 221  
Db 869 CTCATAGAAGATGATGACACTGTCATTGGAAAGTGATTACGCGAGTCTCTGGGACACAATG 928  
Qy 222 GlnValGluLeuProProLeuGluIleAsnSerArgValSerLeuLysGly-GlyGluTh 241  
Db 929 CAGTGTGNACTTCCTCTCTTGGAAATAAATCCAGAGTTTCTTTGAA-GGTTGGAGAAGC 987  
Qy 241 rIleGluSerGlyThrValIlePheCysAspValLeuProGlyLysGluSerLeuGlyTy 261  
Db 988 AATAGAATCTGGAACAGTTATATTCTGTGATGTTTGTCCAGGAAAGAAAGCTTAGGATA 1047  
Qy 261 rPheValGlyValAspMetAspAsnProIleGlyAsnTrpAspGlyArgPheAspGlyVa 281  
Db 1048 TTTTGTGTTGGTGTGGACATGGATAACCTTATTGGCAACTGGGATGGAGATTGTATGGAGT 1107  
Qy 281 1---LeuCysSerPheAlaCysValGluSerThrIleLeuLeuHisIleAsnAspIleI 300  
Db 1108 GCAGCTTTGTAGTTTGTGCGTGTGTGAAAGTACAAATCTATTGGACATCAATGATATCAT 1167  
Qy 300 eProGluSerValThrGlnGluArgProProLysLeuAlaPheMetSerArgGlyVa 320  
Db 1168 CCCAGAGAGTGTGACGAGGAAAGAGAGGCTCCCAAACTTGCGCTTTATGTCAAGAGGTGT 1227  
Qy 320 lGlyAspLysGlySerSerHisAsnLysProLysAlaThrGlySerThrSerAspPr 340  
Db 1228 TGGGACAAAGGTTTCATCCAGTCATAATAAACCAAGGCTACAGGATCTACCTCAGACCC 1287  
Qy 340 oGlyAsnArg---ArgSerGluLeuPheThrThrLeuAsnGlySerSerValAspSerGI 359  
Db 1288 TGGAAATAGAAACAGATCTGAATATTATTTTATACCTTAAATGGTCTCTCTGTGACTACA 1347  
Qy 359 nProGlnSerLysSerLysAsnThrTrpTyrIleAspGluValAlaGluAspProAlaIy 379  
Db 1348 ACCAATCCAAATCAAAAAATACATGTCATGTCATGTCATGTCATGTCATGTCATGTCAT 1407  
Qy 379 sSerLeuThrGluIleSerThrAspPheAspArgSerProProLeuGlnProProPr 399

Db 1408 ATCTCTTACAGAGATATCTACAGACTTTGACCGTTCTTCCACACACTCCAGCCTCCTCC 1467  
Qy 399 oValAsnSerLeuThrThrGluAsnArgPheHisSerLeuProPheSerLeuThrLysMe 419  
Db 1468 TGTGAACCTCACTGACCCACCGAGAACAGATTCCACTCTTTTACCATTCAGTCTCACCAAGAT 1527  
Qy 419 tProAsnThrAsnGlySerIleGlyHisSerProLeuSerLeuSerAlaGlnSerValMe 439  
Db 1528 GCCCAATACCAATGGAAGTATTGGCCACAGTCCACTTTCTCTGTGAGCCAGTCTGTAA 1587  
Qy 439 tGluGluLeuLeuThrAlaProValGlnGluSerProProLeuAlaMetProProGlyAs 459  
Db 1588 GGAAGAGCTAAACACTGCACCCGTCACAGAGAGTCCACCTTGGCCATCGCTCTCTGGGAA 1647  
Qy 459 nSerHisGlyLeuGluValGlySerLeuAlaGluValLysGluAsnProProPheTyrGI 479  
Db 1648 CTCACATGCTCTAGAGTGGGCTCATTTGGCTGAAGTTAAGGAGAACCTCTCTTCTATGG 1707  
Qy 479 yValIleArgTrpIleGlyGlnProProGlyLeuAsnGluValLeuAlaGlyLeuGluLe 499  
Db 1708 GGTAAATCCGTTGGATCGGTGAGCCACACAGGACTGAATGAAGTGTCTGCTGGACTGGAACT 1767  
Qy 499 uGluAspGluCysAlaGlyCysThrAspGlyThrPheArgGlyThrArgTyrPheThrCy 519  
Db 1768 GGAAGATGAGTGTGAGGCTGTACCGATGGAACTTTCAGAGGCACTCGGTATTTCACTCG 1827  
Qy 519 sAlaLeuLysLysAlaLeuPheValLysLeuLysSerCysArgProAspSerArgPheAl 539  
Db 1828 TGCCCTGAAGAGGCGCTGTTTGTGNAACTGAGAGCTGCAGGCTGACTCTAGGTTTGC 1887  
Qy 539 aserLeuGlnProValSerAsnGlnIleGluArgCysAsnSerLeuAlaPheGlyTy 559  
Db 1888 ATCATTTGAGCGCGTTTCCAATCAGATTGAGCGCTGTAATCTTTTAGCATTTGGAGGCTA 1947  
Qy 559 rLeuSerGluValValGluGluAsnThrProProLysMetGluLysGluGlyLeuGluI 579  
Db 1948 CTTAAGTGAAGTAGTAGAAGAAATACTCCACCAAAAATGGAAGAAAGGCTTGGAGAT 2007  
Qy 579 eMetIleGlyLysLysLysGlyIleGlnGlyHisTyrAsnSerCysTyrLeuAspSerPh 599  
Db 2008 AATGATTTGGGAAGAAAGGATCCAGGTCATTACAAATCTTGTACTTAGACTCAAC 2067  
Qy 599 rLeuPheCysLeuPheAlaPheSerSerValLeuAspThrValLeuLeuArgProLysGI 619  
Db 2068 CTTATTTCTCTTATTGCTTTTAGTTCTGTGACACTGTGTACTTAGACCCAAAGA 2127  
Qy 619 uLysAsnAspValGluTyrTyrSerGluThrGlnGluLeuLeuArgThrGluIleValAs 639  
Db 2128 AAAGAACGATGTAGAATATTATAGTGAAACCCAGAGCTACTGAGGACAGAAATTTGTTAA 2187  
Qy 639 nProLeuArgIleTyrGlyTyrValCysAlaThrLysIleMetLysLeuArgLysIle 659  
Db 2188 TCCTCTGAGAATATATGATATGTGTGCCCAAAAAATATGAAACTGAGGAAAAATACT 2247  
Qy 659 uGluLysValGluAlaAlaSerGlyPheThrSerGluGluLysAspProGluGluPheLe 679  
Db 2248 TGAAAGGTGGGCGTGCATCAGGATTTACCTCTGAGAAAAGAGATCTCTGAGGAATCTT 2307  
Qy 679 uAsnIleLeuPheHisIleLeuArgValGluProLeuLeuLysIleArgSerAlaGI 699  
Db 2308 GAATATTCTGTTTCATCATATTTTAAAGGTAGAACCTTTTGTCTTAAATAAAGATCAGCAG 2367  
Qy 699 yGlnLysValGlnAspCysTyrPheThrGlnIlePheMetGluLysAsnGluLysValGI 719  
Db 2368 TCAAAAGGTCAAGATTCCTTCTATCAAAATTTTATGGAATAAAAAATGAGAAAGTTGG 2427  
Qy 719 yValProThrIleGlnGlnLeuLeuGluTyrPheIleAsnSerAsnLeuLysPheAl 739  
Db 2428 CGTTCACCAATTCAGCAGTGTGTAAGATGGTCTTTTATCAACAGTACCTGGAATTTGC 2487  
Qy 739 aGluAlaProSerCysLeuIleIleGlnMetProArgPheGlyLysAspPheLysLeuPh 759  
Db 2488 AGAGGCACCATCATGCTGATTATTTCAGATGCTCGATTGGAAAGACTTTTAAACTATT 2547



QY 140 LysPheProGlyValValArgPheArgGlyProLeuLeuAlaGluArgThrValSerGly 159  
DB 812 AAATTCCTGGAGTTGTACGCTTCAGAGCACCCCTGTTAGCAGAGACAGCTCTCCGGA 871  
QY 160 IlePhePheGlyValGluLeuLeuGluGluArgGlyGlnGlnPheThrAspGlyVal 179  
DB 872 ATATTCCTGGAGTTGAATTCGCGAAGAGGTCGTGTCAGAGTTTCACTGACGGGTG 931  
QY 180 TyrGlnGlyLysGlnLeuPheGlnCysAspGluAspCysGly- --PheValAlaLeuAsp 198  
DB 932 TACCAGGGAACAGCTTTTCAGTGTGATGAAGATTGGCGCTGTTGTTGCATTGGAC 991  
QY 199 LysLeuGluLeuIleGluAspAspThrAlaLeuGluSerAspTyrAlaGlyProGly 218  
DB 992 AAGCTAGAACTCATAGAAGATGACACTGCTGCTTTGGAAATGAATTCAGCAGGTCCTGGG 1051  
QY 219 AspThrMetGlnValGluLeuProProLeuGluIleAsnSerArgValSerLeuLysGly 238  
DB 1052 GACCAATGCGAGGTCGAACTTCCTCCTTTGGAAATGAATTCAGCAGGTTTCTTTGAA- GGT 1110  
QY 239 -GlyGlnThrIleGluSerGlyThrValIlePheCysAspValLeuProGlyLysGluSe 258  
DB 1111 TGGAGAAACAATAGAAATCTGGAAACAGTTATATCTGTGATGTTTGGCCAGGAAAGAAAG 1170  
QY 258 rLeuGlyTyrPheValGlyValAspMetAspAsnProIleGlyAsnTrpAspGlyArgPh 278  
DB 1171 CTTAGGATATTTGTTGTTGGATGGACATGGATTAACCTTATTTGGCACTGGGATGGAGATT 1230  
QY 278 eAspGlyVal- --LeuCysSerPheAlaCysValGluSerThrIleLeuLeuHisIleAs 297  
DB 1231 TCATGGAGTCAGGCTTTCTAGTTTTCGCGTGTGTTGAAAGTACAAATCTATTGACATCAA 1290  
QY 297 nAspIlePro- ----GluSerValThrGlnGluArgArgProProLysLeuAl 314  
DB 1291 TGATATCATCCAGCTTTATCAGAGAGTGTGACGAGGAAAGAGGCGCTCCCAAACTTGC 1350  
QY 314 aPheMetSerArgGlyValGlyAspLysGlySerSerHisAsnLysProLysAlaIath 334  
DB 1351 CTTTATGTCAAGAGTGTGGGACAAAGTTTCATCCAGTCAATATAAACCAAGGCTAC 1410  
QY 334 rGlySerThrSerAspProGlyAsnArg- --ArgSerGluLeuPheTyrThrLeuAsnGl 353  
DB 1411 AGGATCTACCTCAGACCCCTGGAAATAGAAACAGATCTGAATTTATATACCTTAAATGG 1470  
QY 353 ySerSerValAspSerGlnProGlnSerLysSerLysAsnThrTrpTyrIleAspGluVa 373  
DB 1471 GTCTTCTGTGACTCACAAACCAATCCAAATCAAAAAATACATGGTACATTTGATGAAGT 1530  
QY 373 lAlaGluAspProAlaLysSerLeuThrGluIleSerThrAspPheAspArgSerSerPr 393  
DB 1531 TGCAGAACCCCTGCAGAAATCTCTACAGAGATATCTACAGACTTTGACCGTCTTCACC 1590  
QY 393 oProLeuGlnProProValAsnSerLeuThrThrGluAsnArgPheHisSerLeuPr 413  
DB 1591 ACCACTCAGCCTCCTCCTGTGAACTCACTGACCAACGAGAACAGATTCACCTCTTACC 1650  
QY 413 oPheSerLeuThrLysMetProAsnThrAsnGlySerIleGlyHisSerProLeuSerIle 433  
DB 1651 ATTCACTCTACCAAGATGCCAATACCAATGGAAGTATTGGCCACAGCTCCACTTCTCT 1710  
QY 433 uSerAlaGlnSerValMetGluLeuLeuAsnThrAlaProValGlnGluSerProProle 453  
DB 1711 GTCAGCCCGAGTCTGTAATGGGAAGAGCTAAACACTGACCCGTCGAAGAGAGTCCACCTT 1770  
QY 453 uAlaMetProProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLysGl 473  
DB 1771 GGCCATGCTCTCTGGGAACCTACATGGTCTAGAGTGGGCTCATTTGGCTGAAGTTAAGGA 1830  
QY 473 uAsnProProPheTyrGlyValIleArgTrpIleGlyGlnProProGlyLeuAsnGluVa 493  
DB 1831 GAACCCCTCTTCTATGGGGTAAATCCGTTGGATCGTTCAGCCACGAGGACTGAATGAAGT 1890

QY 493 lLeuAlaGlyLeuGluLeuGluAspGluCysAlaGlyCysThrAspGlyThrPheArgGl 513  
DB 1891 GCTCGCTGGACTGGAACTGGAAGATGAGTGTGAGGCTGTACGATGGAACCTTCAGAGG 1950  
QY 513 yThrArgTyrPheThrCysAlaLeuLysLysAlaLeuPheValLysLeuLysSerCysAr 533  
DB 1951 CACTCGGTATTTACCTGTGCCCCTGAAGAAGCGCTGTTGTGAAACTGAAGAGCTGCAG 2010  
QY 533 gProAspSerArgPheAlaSerLeuGlnProValSerAsnGlnIleGluArgCysAsnSe 553  
DB 2011 GCCTGACTCTAGTTTGCATCATTCGAGCCGGTTTCCAAATCAGATTGAGCGCTGAATCT 2070  
QY 553 rLeuAlaPheGlyTyrLeuSerGluValValGluGluAsnThrProProLysMetGl 573  
DB 2071 TTTAGCATTTGGAGGCTACTTAAGTGAAGTAGTAGAAGAAATACTCCACCAAAATGGA 2130  
QY 573 uLysGluGlyLeuGluIleMetIleGlyLysLysGlyIleGlnGlyHisTyrAsnSe 593  
DB 2131 AAAAGAGAGCTTTGGAGATAATGATTGGGAAGAAGAGCATCCAGGGGTCAATTACAATTC 2190  
QY 593 rCysTyrLeuAspSerThrLeuPheCysLeuPheAlaPheSerSerValLeuAspThrVa 613  
DB 2191 TTGTTACTTAGACTCAACTTATCTGCTTATTTGCTTTTGTCTGTCGACTGT 2250  
QY 613 lLeuLeuArgProLysGluLysAsnAspValGluTyrTyrSerGluThrGlnGluLeuLe 633  
DB 2251 GTTACTTAGACCCCAAGAAAGAACGATGTAGAATAATTATAGTGAACCCCAAGAGCTACT 2310  
QY 633 uArgThrGluIleValAsnProLeuArgIleTyrGlyTyrValCysAlaThrLysIleMe 653  
DB 2311 GAGGACAGAAATTTGTTAATCTCTCAGAAATATATGATATGTGTGTCACCAAAATTA 2370  
QY 653 tLysLeuArgLysIleLeuGluLysValGluAlaAlaSerGlyPheThrSerGluGluLyl 673  
DB 2371 GAAACTGAGGAAATACTTTGAAAGGTGGAGCTGCATCAGGATTTACCTCTCGAAGAAA 2430  
QY 673 sAspProGluGluPheLeuAsnIleLeuPheHisIleLeuArgValGluProLeuLe 693  
DB 2431 AGATCCTGAGGAATTTCTGAATATTTCTGTTTCATCATATTTTAAGGTAGAACCTTGT 2490  
QY 693 uLysIleArgSerAlaGlyGlnLysValGlnAspCysTyrPheTyrGlnIlePheMetGl 713  
DB 2491 AAAAATAAGATCAGCAGGTCAAAAGGTACAAAGATTGTTACTTCTATCAAAATTTTATGGA 2550  
QY 713 uLysAsnGluLysValGlyValProThrIleGlnGlnLeuLeuGluTrpSerPheIleAs 733  
DB 2551 AAAAATGAGAAAGTTGGCGTTCCCACAATTCAGCAGTTGTTAGAAATGGTCTTTTATCAA 2610  
QY 733 nSerAsnLeuLysPheAlaGluAlaProSerCysLeuIleIleGlnMetProArgPheGl 753  
DB 2611 CAGTAACCTGAATTTGACAGGCGACCATCATGCTGTGATTATTCAGATGCTCGATTGG 2670  
QY 753 yLysAspPheLysLeuPheLysIlePhePheProSerLeuGluLeuAsnIleThrAspLe 773  
DB 2671 AAAAGACTTTAAACTATTAAAAAAATTTTCTCTCTCGAAATTTAAATATAACAGATTT 2730  
QY 773 uLeuGluAspThrProArgGlnCysArgIleCysGlyGlyLeuAlaMetTyrGluCysAr 793  
DB 2731 ACTTGAAGACACTCCCGACAGTCGCGATATGTGGAGGCTTGAATGTATGAGTGTAG 2790  
QY 793 gGluCysTyrAspAspProAspIleSerAlaGlyLysIleLysGlnPheCysLysThrCy 813  
DB 2791 AGAATGCTACGACGATCCGGACATCTCAGCTGGAATAATCAAGCAGTTTGTAAAAACCTG 2850  
QY 813 sAsnThrGlnValHisLeuHisProLysArgLeuAsnHisLysTyrAsnProValSerLe 833  
DB 2851 CAACACTCAAGTCCACTTCATCCGAGAGGCTGAATCATATAATATAACCCAGTGTCACT 2910  
QY 833 uProLysAspLeuProAspTrpAspTrpArgHisGlyCysIleProCysGlnAsnMetGl 853  
DB 2911 TCCCAAGAGACTTACCCGACTGGGACTGGAGACACGGCTGCATCCCTTGCAGAAATATGGA 2970  
QY 853 uLeuPheAlaValLeuCysIleGluThrSerHisTyrValAlaPheValLysTyrGlyLy 873







356 lAspSerGlnProGlnSerLysSerLysAsnThrTrpTyrIleAspGluValAlaGluAs 376  
Db TGACTCAACCAACCAATCCAAATCAAAATACATGTTACATGATGAAGTTGCAGAGA 1381  
376 pProAlaLysSerLeuThrGluIleSerThrAspPheAspArgSerSerProLeuG1 396  
Db CCGTGCAGAAATCTCTTACAGAGATATCTACAGACTTTGACCGTCTCTTACCACCACTCCA 1441  
396 nProProValAsnSerLeuThrThrGluAsnArgPheHisSerLeuProPheSerLe 416  
Db GCCTCTCTCTGTGAACCTCACTGACCAACCGAGAACAGATTCACCTTTACCAATCAGTCT 1501  
416 uThrLysMetProAsnThrAsnGlySerIleGlyHisSerProLeuSerLeuAlaG1 436  
Db CACCAAGATGCCAATACCAATGGAAGTATTGGCCACAGTCCACTTTCTCTGTGACGCCA 1561  
436 nSerValMetGluLeuLeuAsnThrAlaProValGlnGluSerProProLeuAlaMetPr 456  
Db GTCTGTATGAAGAGCTAAACACTGCACCCCGTCCAGAGAGTCCACCCCTTGGCCATGCC 1621  
456 oProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLysGluAsnProPr 476  
Db TCCTGGGAACTCAATGCTTAGAAGTGGGCTCATTTGGCTGAAGTTAAGGAGAACCTTCC 1681  
476 oPheTyrGlyValIleArgTrpIleGlyGlnProProGlyLeuAsnGluValLeuAlaG1 496  
Db TTTCTATGGGTAATCCGTTGGATCGGTGAGCCAGCCAGGACTGATGAAGTGTCTGCTGG 1741  
496 YLeuGluLeuGluAspGluCysAlaGlyCysThrAspGlyThrPheArgGlyThrArgTy 516  
Db ACTGGAATGGAAGATGATGTGAGGCTGTGAGGCTGACCGATGAACTTCAGAGGCACTCG 1801  
516 rPheThrCysAlaLeuLysLysAlaLeuPheValLysLysLysSerCysAtgProAspSe 536  
Db TTTTACCTGTGCCCTGAAGAGCGCTGTGTGAACTGGAAGACTGAGAGCTGAGGCTGACTC 1861  
536 rArgPheAlaSerLeuGlnProValSerAsnGlnIleGluArgCysAsnSerLeuAlaPh 556  
Db TAGTTTGCATATTGCAGCGGTTTCCAAATCAGATTGAGCGCTGTAATCTTTAGCAAT 1921  
556 eGlyGlyTyrLeuSerGluValValGluGluAsnThrProProLysMetGluLysGluG1 576  
Db TGGAGGCTACTTAAGTGAAGTAGTAGAAGAAATATCTCCACCAAAATGGAAGAAAGG 1981  
576 YLeuGluIleMetIleGlyLysLysLysGlyIleGlnGlyHisTyrAsnSerCysTyrIle 596  
Db CTGGAGATAATGATTGGGAAGAAAGGCAATCCAGGCTCATTAACAATCTTTGTACTT 2041  
596 uAspSerThrLeuPheCysLeuPheAlaPheSerSerValLeuAspThrValLeuLeuAr 616  
Db AGACTCAACCTTATCTGCTTATTTGCTTTTATGTTCTGTTGAGACATGTTGTTACTTAG 2101  
616 gProLysGluLysAsnAspValGluTyrTyrSerGluThrGlnGluLeuLeuArgThrG1 636  
Db ACCCAAGAAAGAAACGATGAGATAATATATAGTGAACCCCAAGAGCTACTGAGGACAGA 2161  
636 uIleValAsnProLeuArgIleGlyTyrValCysAlaThrLysIleMetLysLeuAr 656  
Db AATTGTTAATCTCTGAGAAATATGAGATATGATGTGTGCCACAAAAATATGAAACTGAA 2221  
656 g-LysIleLeuGluLysValGluAlaAlaSerGlyPheThrSerGluGluLysAspProG 676  
Db -GAAATATCTTGAAGAGTGGAGGCTGCATCAGGATTTACCTCTGAGGAAAGAAAGTCTCG 2280  
676 luGluPheLeuAsnIleLeuPheHisIleLeuArgValGluProLeuLeuLysIleA 696  
Db AGGAATCTTGAATATTCTGTTTCATCATATTATTAAGGGTAGAACCTTTTGTGTAATAA 2340  
696 rGSerAlaGlyGlnLysValGlnAspCysTyrPheTyrGlnIlePheMetGluLysAsnG 716  
Db GATCAGCAGGTCAAAAGGTACAGATTTGTTACTTCTATCAAAATTTTATGGAAAAAATG 2400  
716 luLysValGlyValProThrIleGlnGlnLeuLeuGluTrpSerPheIleAsnSerAsnL 736

2401 AGAAAGTTGGCGTTCCCAATTCAGCAGTTGTTAGAAATGGTCTTTTATCAACAGTAACC 2460  
Qy euLysPheAlaGluAlaProSerCysLeuLeuIleGlnMetProArgPheGlyLysAspP 756  
Db TGAATTTGACAGGACACATCATGCTTGATTTATTCAGATGCTCGATTTGGAAGACT 2520  
756 heLysLeuPheLysLysIlePheProSerLeuGluLeuAsnIleThrAspLeuLeuGluA 776  
Db TTAACATTTTAAAAAATTTTCTCTCTGGAATTTAAATATAACAGATTTTACTTTGAAG 2580  
776 spThrProArgGlnCysArgIleCysGlyLeuAlaMetTyrGluCysArgGluCysT 796  
Db ACACCTCCAGACAGTCCGCGATATGTGAGGCGCTTGCATGTATGATGATGATGATGATG 2640  
796 rAspAspProAspIleSerAlaGlyLysIleLysGlnPheCysLysThrCysAsnThrG 816  
Db ACGAGATCCGAGCATCTCAGCTGGAANAATCAGACGTTTTGTAAACCTGCAACACTC 2700  
816 InValHisLeuHisProLysArgLeuAsnHisLysTyrAsnProValSerLeuProLysA 836  
Db AAGTCCACCTTCATCCGAAGAGGCTGAATCATATAATATAACCCAGTGTCACTTCCCAAAG 2760  
836 spLeuProAspTrpAspTrpArgHisGlyCysIleProCysGlnAsnMetGluLeuPheA 856  
Db ACTTACCCGACTGGGACTGGAGACACGCTGCATCCCTTGCAGAAATATGGAGTTATTG 2820  
856 laValLeuCysIleGluThrSerHisTyrValAlaPheValLysTyrGlyLysAspAsP 876  
Db CTGTTCTCTGCATAGAAACAGCCACTATGTTGCTTTTGTGAAGTATGGGAAGGAGATT 2880  
876 eAlaTrpLeuPhePheAspSerMetAlaAspArgAspGlyGlyGlnAsnGlyPheAsnI 896  
Db CTGCTCTGCTCTCTTTGACAGCATGGCGATGGGATGGTGTGATGATGCTTCAACA 2940  
896 leProGlnValThrProCysProGluValGlyGluTyrIleLysMetSerLeuGluAspL 916  
Db TTTCTCAAGTCAACCCCATGCCAGAAAGTAGGAGAGTACTTGAAGATGTCTCTGGAAGAC 3000  
916 euHisSerLeuAspSerArgIleGlnGlyCysAlaArgLeuLeuCysAspAlat 936  
Db TGCATTTCTTGGACTCCAGGAGAAATCCAAAGGCTGTGCAAGAGCTGCTTTGTGATGCAT 3060  
936 yrMetCysMetTyrGlnSerProThrMetSerLeuTyrLys 949  
Db ATATGTGATGTACCAGAGTCCACATGAGTTGTACAAA 3101  
RESULT 7  
LOCUS AR338799 2523 bp DNA linear PAT 17-AUG-2003  
DEFINITION Sequence 290 from patent US 6569662.  
ACCESSION AR338799  
VERSION AR338799.1 GI:33725656  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 2523)  
AUTHORS Tang, Y. T., Zhou, P. and Drmanac, R. T.  
TITLE Nucleic acids and polypeptides  
JOURNAL Patent: US 6569662-A 290 27-MAY-2003;  
FEATURES Location/Qualifiers  
          1..2523  
          /organism="unknown"  
          /mol\_type="genomic DNA"  
ORIGIN  
Alignment Scores:  
Pred. No.: 0 Length: 2523  
Score: 641.00 Matches: 668  
Percent Similarity: 99.55% Conservative: 0  
Best Local Similarity: 99.55% Mismatches: 0  
Query Match: 67.54% Indels: 3

DB:	6	Gaps:	0
US-09-671-687A-3 (1-949) x AR338799 (1-2523)			
QY	282	LeuCysSerPheAlaCysValGluSerThrIleLeuLeuHisIleAsnAspIleIlePro	301
Db	129	CTTTGTAGTTTGGCTGTGTGAAAGTACAAATCTATTGCACATCAATCATATCATCCCA	198
QY	302	GluSerValThrGlnGluArgArgProProLysLeuAlaPheMetSerArgGlyValGly	321
Db	189	GAGAGTGAGCCAGAGAAAGAGGCTCCCAACTTGCCTTTATGTCAAGAGGTGTGG	248
QY	322	AspLysGlySerSerHisAsnLysProLysAlaThrGlySerThrSerAspProGly	341
Db	249	GACAAAGGTTCATCCAGTCATAATAAACCAAGGCTACAGGATCTACCTCAGACCTGGA	308
QY	342	AsnArg---ArgSerGluLeuPheThrLeuAsnGlySerSerValAspSerGlnPro	360
Db	309	AATAGAAAACAGATCTGAATATTATTTATACCTTTAAATGGGTCTTCTGTGACTCACAA	368
QY	361	GlnSerLysSerLysAsnThrTrpTyrIleAspGluValAlaGluAspProAlaLysSer	380
Db	369	CAATCCAAATCAAAATAACATGGTACATGTGATGAAGTTGCAGAACCTCGAAATCT	428
QY	381	LeuThrGluIleSerThrAspPheAspArgSerSerProLeuGlnProProVal	400
Db	429	CTTACAGAGATATCTACAGACTTTGACCGTTCTTCCACCACTCCAGGCTCCTCTGTG	488
QY	401	AsnSerLeuThrThrGluAsnArgPheHisSerLeuProPheSerLeuThrLysMetPro	420
Db	489	AACTCACTGACCCAGACAGATTCACACTCTTTACCATTTCAGTCTCACCAGATGCC	548
QY	421	AsnThrAsnGlySerIleGlyHisSerProLeuSerLeuSerAlaGlnSerValMetGlu	440
Db	549	AATACCAATGGAAGTATTGGCCACAGTCCACTTCTCTGACCCAGCTGTGTAATGGA	608
QY	441	GluLeuAsnThrAlaProValGlnGluSerProProLeuAlaMetProProGlyAsnSer	460
Db	609	GAGCTAACACTGCACCCGCTCCAGAGAGTCCACCCTTGGCCATGGCTCTCGGAACTCA	668
QY	461	HisGlyLeuGluValGlySerLeuAlaGluValLysGluAsnProProPheThrGlyVal	480
Db	669	CATGGCTAGAAAGTGGGCTCATTTGGCTGAAGTTAAGAGAACCCCTCTTCTATGGGTA	728
QY	481	IleArgTrpIleGlyGlnProProGlyLeuAsnGluValLeuAlaGlyLeuGluLeuGlu	500
Db	729	ATCCGTTGGATCGGTGAGCCAGCACAGACTGAATGAAGTCTCGCTGGAACCTGGAA	788
QY	501	AspGluCysAlaGlyCysThrAspGlyThrPheArgGlyThrArgGlyPheThrCysAla	520
Db	789	GATGAGTGTGAGGCTGTACGGATGGAACCTTCAGAGGCACCTCGGTATTTTCACTCTGCC	848
QY	521	LeuLysLysAlaLeuPheValLysLeuLysSerCysArgProAspSerArgPheAlaSer	540
Db	849	CTGAAGAGGCGCTCTTTGTGAACTGAAGAGCTGCAGCGCTGACTCTAGGTTTGATCA	908
QY	541	LeuGlnProValSerAsnGlnIleGluArgCysAsnSerLeuAlaPheGlyGlyThrLeu	560
Db	909	TTGCAGCCGGTTTCCAAATCAGATTGAGCGCTGTAACCTTTAGCATTTGGAGGCTACTTA	968
QY	561	SerGluValValGluGluAsnThrProProLysMetGluLysGluGlyLeuGluIleMet	580
Db	969	AGTGAAGTAGTAGAAGAAATACTCCACCAAAATGGAAGAAAGAGGCTTGGAGATAATG	1028
QY	581	IleGlyLysLysGlyIleGlnGlyHisThrAsnSerCysThrLeuAspSerThrLeu	600
Db	1029	ATTGGAGAGAGAAAGGATCCAGGTCATTACATCTTGTACTTAGACTCAACTTA	1088
QY	601	PheCysLeuPheAlaPheSerSerValLeuAspThrValLeuLeuArgProLysGluLys	620
Db	1089	TTCTGCTATTGCTTTAGTTCTGTCTGGACACTGTGTACTTAGACCCAAAGAAAG	1148
QY	621	AsnAspValGluTyrThrSerGluThrGlnGluLeuLeuArgThrGluIleValAsnPro	640

## RESULT 8

CQ719792

LOCUS

DEFINITION

ACCESSION

VERSION

Sequence 5726 from Patent WO02068579.

CQ719792

CQ719792.1

GI:42280649

3480 bp

DNA

linear

PAT 03-FEB-2004

## KEYWORDS

SOURCE Homo sapiens (human)

## ORGANISM

Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

## REFERENCE

1 Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.

AUTHORS  
TITLE  
Kits, such as nucleic acid arrays, comprising a majority of  
humanexons or transcripts, for detecting expression and other uses  
thereof

JOURNAL Patent: WO 02068579-A 5726 06-SEP-2002;

PE Corporation (US)

## FEATURES

Location/Qualifiers

source

1..3480

/organism="Homo sapiens"

/mol\_type="unassigned DNA"

/db\_xref="taxon:9606"

## ORIGIN

Alignment Scores:

Pred. No.:	0	Length:	3480
Score:	640.00	Matches:	947
Percent Similarity:	98.24%	Conservative:	0
Best Local Similarity:	98.24%	Mismatches:	2
Query Match:	67.44%	Indels:	17
DB:	6	Gaps:	0

US-09-671-687A-3 (1-949) x CQ719792 (1-3480)

Qy 1 MetSerGlyLeuTrpSerGlnGluLysValThrSerProTyrTrpGluArgIle 20  
 Db 235 ATGAGTTTCAGGCTTATGAGGCAAGAAAGATCACTTCACCTACTGGGAGCGGATT 294  
 Qy 21 PheTyrLeuLeuGlnGluCysSerValThrAspLysGlnThrGlnLysLeuLys 40  
 Db 295 TTTTACTTGTCTTCAAGATGTCAGCGTTACAGACAAACAAACAAAGCTCTTAA 354  
 Qy 41 ValProLysGlySerIleGlyGlnTyrIleGlnAspArgSerValGlyHisSerArgIle 60  
 Db 355 GTACCGAAGGAGATATAGACAGATATATTCAGATCGTTCTGGGGCATTCAGGATT 414  
 Qy 61 ProSerAlaLysGlyLysLysAsnGlnIleGlyLeuLysIleLeuGluGlnProHisala 80  
 Db 415 CCTTCTGCAAAAGCAAGCAAAATTCAGATTGGATTAAAAATTTAGAGCAACCTATGCA 474  
 Qy 81 ValLeuPheValAspGlu---AspValValGluIleAsnGluLysPheThrGluLeuLeu 99  
 Db 475 GTTCTCTTTGTGATGAAAGGATGTTGTAGAGATAAAATGAAAGTTTCACAGAGTTACTT 534  
 Qy 100 LeuAlaIleThrAsnCysGluGluArgPheSerLeuPheLysAsnArgAsnArgLeuSer 119  
 Db 535 TTGGCAATTCAATTTGAGAGAGGTTTACCGCTGTTTAAAGATACAGATCTGGGAGAA 594  
 Qy 120 LysGlyLeuGlnIleAspValGlyCysProValLysValGlnLeuArgSerGlyGluGlu 139  
 Db 595 AAAGGCTCCAAATAGACGTGGCTGCTCTGTGAAAGTACAGCTCAGATCTGGGAGAA 654  
 Qy 140 LysPheProGlyValValArgPheArgGlyProLeuLeuAlaGluArgThrValSerGly 159  
 Db 655 AAATTTCTGGAGTTGTACGTTTCAGAGCACCCCTGTTAGCAGAGAGGACAGCTCCGGA 714  
 Qy 160 IlePhePheGlyValGluLeuLeuGluGluArgGlyArgGlyGlnGlyPheThrAspGlyVal 179  
 Db 715 ATATTCTTTGGAGTTGAATTCGTGAAAGAGTCTGTGTCAAGGTTTCACTACCGGGTG 774  
 Qy 180 TyrGlnGlyLysGlnLeuPheGlnCysAspGluAspCysGly---PheValAlaLeuAsp 198  
 Db 775 TACCAGGGAACACGCTTTTCAGTGTGATGAAGATTGTGGCGTGTGTTCATTTGGAC 834  
 Qy 199 LysLeuGluLeuIleGluAspAspThrAlaLeuGluSerAspTyrAlaGlyProGly 218  
 Db 835 AAGCTAGAACTCATAGAAGATGATGACACTGCTATTGGAAAGTGATTACGACGCTCTGGG 894

Qy 219 AspThrMetGlnValGluLeuProProLeuGluIleAsnSerArgValSerLeuLysGly 238  
 Db 895 GACACAATGCAAGTCCAACTTCTCTCTTTGAAATAAAATCCAGAGTTCTTTTGAAG- 953  
 Qy 239 -GlyGluThrIleGluSerGlyThrValIlePheCysAspValLeuProGlyLysGlu 258  
 Db 954 TGGAGAAACAATAGAAATCTGGAACAGTTATATTTCTGTGATGTTTGGCCAGAAAGAAAG 1013  
 Qy 258 rLeuGly-TyrPheValGlyValAspMetAspAsnProIleGlyAsnTrpAspGlyArgP 278  
 Db 1014 CTTACG-ATATTTTGTGTGTGGTGGACATGGATAACCTATTGGCAACTGGGATGAAGAT 1072  
 Qy 278 heAspGlyVal---LeuCysSerPheAlaCysValGluSerThrIleLeuLeuHisIleA 297  
 Db 1073 TTGATGGAGTGCAGCTTTGTAGTTTGGCTGTGTGAAAGTACAATTTCTATTGCACATCA 1132  
 Qy 297 snAspIleIleProGluSerValThrGlnGluArgProProLysLeuAlaPheMetS 317  
 Db 1133 ATGATATCATCCAGAGAGTGTGACGAGGAAAGAGGCGCTCCCAAACTTGCTTTATGT 1192  
 Qy 317 erArgGlyValGlyAspLysGlySerSerHisAsnLysProLysAlaThrGlySerT 337  
 Db 1193 CAAGAGGTGTGGGACAAAGTTTCATCCAGTCATAATAAACCAAGGCTACAGGATCTA 1252  
 Qy 337 hrSerAspProGlyAsnArg---ArgSerGluLeuPheTyrThrLeuAsnGlySerSerV 356  
 Db 1253 CCTCAGACCTCGAAATAGAAACAGATCTGAATTTATTTATACCTTAATGGGTCTTCG 1312  
 Qy 356 alAspSerGlnProGlnSerLysSerLysAsnThrTrpTyrIleAspGlu-ValAlaGlu 375  
 Db 1313 TTGACTCACACCAATCCAAATCAAAATAATACATGTTACATGATGAAGGTGTCAGAA 1372  
 Qy 376 aspProAlaLysSerLeuThrGluIleSerThrAspPheAspArgSerSerProProLeu 395  
 Db 1373 GACCTGCAAAATCTCTTACAGAGATATCTACAGACTTTTGACCGTCTTCCACCACTC 1432  
 Qy 396 GlnProProValAsnSerLeuThrThrGluAsnArgPheHisSerLeuProPheSer 415  
 Db 1433 CAGCCTCTCTCTGTAACCTCACTGACCCAGGAGACAGATTCACCTCTTTTACCATTCAGT 1492  
 Qy 416 LeuThrLysMetProAsnThrAsnGlySerIleGlyHisSerProLeuSerLeuSerAla 435  
 Db 1493 CTCACCAAGATGCCCAATACCAATGGAAGTATTTGGCCACAGTCCACTTTCTCTGTGAGCC 1552  
 Qy 436 GlnSerValMetGluGluLeuAsnThrAlaProValGlnGluSerProProLeuAlaMet 455  
 Db 1553 CAGTCTGTAATGGAAGAGCTAAACACTGCACCCGTCGAAGAGAGTCCACCTTTGGCCATG 1612  
 Qy 456 ProProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLysGluAsnPro 475  
 Db 1613 CCTCTCTGGAACTCAGTGGTCTAGAGTGGGCTCATTTGGCTGAGTTAAGGAGAACCTT 1672  
 Qy 476 PropheTyrGlyValIleArgTrpIleGlyGlnProProGlyLeuAsnGluValLeuAla 495  
 Db 1673 CTTTCTATGGGTAATCGTTGGATCGGTGAGCCACAGGACTGAATGAAGTGTCTCGT 1732  
 Qy 496 GlyLeuGluLeuGluAspGluCysAlaGlyCysThrAspGlyThrPheArgGlyThrArg 515  
 Db 1733 GGACTGGAACCTGGAAGATGAGTGTGCGAGCTGTACGATGGAACCTTCAGAGGACATCGG 1792  
 Qy 516 TyrPheThrCysAlaLeuLysLysAlaLeuPheValLysLeuLysSerCysArgProAsp 535  
 Db 1793 TATTTTCACTGTGCTTGAAGAGGCGCTGTTTGTGAAACTGAAGAGCTGCGGCCCTGAC 1852  
 Qy 536 SerArgPheAlaSerLeuGlnProValSerAsnGlnIleGluArgCysAsnSerLeuAla 555  
 Db 1853 TCTAGTTTGCATCTTCAGCCGCTTTCATCATGATTCAGATTCAGGCTGTAACCTTTAGCA 1912  
 Qy 556 PheGlyGlyTyrLeuSerGluValValGluGluAsnThrProProLysMetGluLysGlu 575  
 Db 1913 TTTGGAGGCTACTTAAGTGAAGTAGAAGAAATACTCCACCAAAATGGAAGAAAGAA 1972  
 Qy 576 GlyLeuGluIleMetIleGlyLysLysGlyIleGlnGlyHisTyrAsnSerCysTyr 595

```
Db 1973 GGCCTGGAGATAATGATTGGGAGAGAAAGGCATCCAGGTCATTACAAATTCCTGTTAC 2032
Qy 596 LeuAspSerThrLeuPheCysLeuPheAlaPheSerSerValLeuAspThrValLeuLeu 615
Db 2033 TTAGACTCAACCTTATCTTCCTTATTGCTTTTGTGTTCTGACACACTGTTACTT 2092
Qy 616 ArgProLysGluLysAsnAspValGluTyrTyrSerGluThrGlnGluLeuLeuArgThr 635
Db 2093 AGACCCAAAGAAAGAACCATGTAGATATTATAGTGAACCCAGAGCTACTGAGGACA 2152
Qy 636 GluLeuValAsnProLeuArgIleTyrGlyTyrValCysAlaThrLysIleMetLysLeu 655
Db 2153 GAAATGTTAATCTCTGAGATATATGATATGTGTGTGCCACAAAATTTATGAAACTG 2212
Qy 656 ArgLysIleLeuGluLysValGluAlaAlaSerGlyPheThrSerGluGluLysAspPro 675
Db 2213 AGGAAATACTTTGAAAGGTGGAGGCTGCATCAGGATTTACCTCTGAAGAAAAGATCCT 2272
Qy 676 GluGluPheLeuAsnIleLeuPheHisIleLeuArgValGluProLeuLeuLysIle 695
Db 2273 GAGGAATCTTGAATATCTGTTTCATCATATTTTAAGGTAGAACCTTTGCTTAAATAA 2332
Qy 696 ArgSerAlaGlyGlnLysValGlnAspCysTyrPheTyrGlnIlePheMetGluLysAsn 715
Db 2333 AGATCAGCAGGTCAAAGGTACAGATGTTACTTCTATCAATTTTATGGAATAAAT 2392
Qy 716 GluLysValGlyValProThrIleGlnGlnLeuLeuGluTyrSerPheIleAsnSerAsn 735
Db 2393 GAGAAAGTTGGCGTTCCCAACAATTCAGCAGTTGTTAGAAATGGTCTTTTATCAACAG 2452
Qy 736 LeuLysPheAlaGluAlaProSerCysLeuIleIleGlnMetProArgPheGlyLysAsp 755
Db 2453 CTGAAATTTGCAGAGCACCATCATGCTCGATTATTCAGATGCTCGATTGGAAGAAG 2512
Qy 756 PheLysLeuPheLysLysIlePheProSerLeuGluLeuAsnIleThrAspLeuLeuGlu 775
Db 2513 TTTAACACTTTTAAAAAATTTTCTCTCTGGAATTTAAATATAACAGATTACTTTGAA 2572
Qy 776 AspThrProArgGlnCysArgIleCysGlyGlyLeuAlaMetTyrGluCysArgGluCys 795
Db 2573 GACACTCCAGACAGTCCCGGATATGTGGAGGCTTGCATATGATGATGAGTGAGAGATGC 2632
Qy 796 TyrAspAspProAspIleSerAlaGlyLysIleLysGlnPheCysLysThrCysAsnThr 815
Db 2633 TACGACGATCCGGACATCTCAGCTCGAAAATCAAGCAGTTTGTGTAACCTGCAACACT 2692
Qy 816 GlnValHisLeuHisProLysArgLeuAsnHisLysTyrAsnProValSerLeuProLys 835
Db 2693 CAAGTCCACCTTCATCCGAGAGGCTGAATCATTAATATAACCCAGTGTCACCTCCCAA 2752
Qy 836 AspLeuProAspTrpAspTrpArgHisGlyCysIleProCysGlnAsnMetGluLeuPhe 855
Db 2753 GACTTACCGCATGGGACTGGAGACACGCGCTGCATCCCTTCCAGAAATATGGAGTTATT 2812
Qy 856 AlaValLeuCysIleGluThrSerHisTyrValAlaPheValLysTyrGlyLysAspAsp 875
Db 2813 GCTGTCTCTCGATAGAACAAAGCCACTATGTTGCTTTTGTGAAGTATGGGAGGACGAT 2872
Qy 876 SerAlaTyrPhePheAspSerMetAlaAspArgAspGlyGlyGlnAsnGlyPheAsn 895
Db 2873 TCTGCTCGCTCTTTTGTACAGCATGGCCGATCGGGATGGGTGAGATGCTCTCAAC 2932
Qy 896 IleProGlnValThrProCysProGluValGlyGluTyrLeuLysMetSerLeuGluAsp 915
Db 2933 ATTCTCAAGTCACCCATGCCAGAGTAGGAGAGTACTTGAAGATGCTCTGGAAGAC 2992
Qy 916 LeuHisSerLeuAspSerArgIleGlnGlyCysAlaArgArgLeuLeuCysAspAla 935
Db 2993 CTGCAATCTCTGAGCTCCAGGAGAAATCCAAGGCTGTGCAGAGACTGCTGTTGTATGCA 3052
Qy 936 TyrMetCysMetTyrGlnSerProThrMetSerLeuTyrLys 949
```

```
Db 3053 TATATGTGCATGTACCAGAGTCCAAACATGAGTTTGTACAAA 3094
RESULT 9
BD160617 2845 bp DNA linear PAT 17-JAN-2003
LOCUS Primer for synthesizing full-length cDNA and use thereof.
DEFINITION BD160617.1 GI:27866375
ACCESSION BD160617
VERSION JP 2002191363-A/15460.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 2845)
AUTHORS Ota T., Isegai T., Nishikawa T., Hayaehi K., Saito K., Yamamoto, J.,
Ishii, S., Sugiyama, T., Wakamatsu, A., Nagai, K. and Otsuki, T.
TITLE Primer for synthesizing full-length cDNA and use thereof
JOURNAL Patent: JP 2002191363-A 15460 09-JUL-2002;
HELIIX RESEARCH INSTITUTE
COMMENT OS Homo sapiens (human)
PN JP 2002191363-A/15460
FD 09-JUL-2002
PF 28-JUL-2000 JP 2000280990
PI TOSHIO OTA, TAKAO ISOGAI, TETSUO NISHIKAWA, KOJI HAYASHI, KAORU
PI SAITO.
PI JUNICHI YAMAMOTO, SHIZUKO ISHII, TOMOYASU SUGIYAMA, AI WAKAMATSU,
PI KEIICHI NAGAI, TETSUJI OTSUKI
PC
C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/ PC
10,
PC C12P21/02, C12Q1/68//C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC
Primer for synthesizing full-length cDNA and use thereof FH Key
Location/Qualifiers (452)..(2644).
FT CDS Location/Qualifiers
source
1..2845
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
```

## ORIGIN

```
Alignment Scores:
Pred. No.: 0 Length: 2845
Score: 608.00 Matches: 861
Percent Similarity: 98.51% Conservative: 0
Best Local Similarity: 98.51% Mismatches: 2
Query Match: 64.07% Indels: 13
DB: 6 Gaps: 0
US-09-671-687A-3 (1-949) x BD160617 (1-2845)
Qy 87 AspValValGluLeuAsnGluLysPheThrGluLeuLeuAlaIleThrAsnCysGlu 106
Db 47 GATGTTGTAGAGATAAATGAAAAGTTTCACAGAGTTACTTTTGGCAATATCCAAATTTGTGAG 106
Qy 107 GluArgPheSerLeuPheLysAsnArgAsnArgSerLysGlyLeuGlnIleAspVal 126
Db 107 GAGAGGTTTCAGCTGTTTAAAAACAGAAAACAGACTAAGTAAAGGCTCCAAATAGACGTG 166
Qy 127 GlyCysProValLysValGlnLeuArgSerGlyGluGluLysPheProGlyValValArg 146
Db 167 GCGTGTCTGTGAAAGTACAGCTGAGATCTGGGGAAGAAAAATTTCTTGAGTGTGTACGC 226
Qy 147 PheArgGlyProLeuLeuAlaGluArgThrValSerGlyIlePheGlyValGluLeu 166
Db 227 TTCAGAGGACCCCTGTTAGCAGAGAGACAGTCTCCGGAATATTTCTTGAGTGTGAATTG 286
Qy 167 LeuGluGluGlyArgGlyGlnGlyPheThrAspGlyValTyrGlnGlyLysGlnLeuPhe 186
Db 287 CTGGAAGAAGAGTCTGCTGCTCAAGGTTTCACTGACGGGGTGTACCAAGGGAACAGCTTTT 346
Qy 187 GlnCysAspGluAspCysGly---PheValAlaLeuAspLysLeuGluLeuIleGluAsp 205
```

Db 347 CAGTGTGANTGAAGATTGTGGCGTGTGTTGTCGATTGGACAAGCTAGAACTCATAGAAGAT 406  
Qy 206 AspAspThrAlaLeuGluSerAspTyrAlaGlyProGlyAspThrMetGlnValGluLeu 225  
Db 407 GATGACACTGATTTGGAAAGTGATTACCGAGTCTCTGGGACACAAATGCGAGTTCGAATT 466  
Qy 226 ProProLeuGluIleAsnSerArgValSerLeuIleGly-GlyGluThrIleGluSerGI 245  
Db 467 CCTCCTTTGGAAATAAATCCAGAGTTCCTTTGAA-GGTTGGAGAAACAATAGAAATCTGG 525  
Qy 245 YThrValIlePheCysAspValLeuProGlyIleGluSerLeuGlyTyrPheValGlyVa 265  
Db 526 AACAGTTATATCTCTGTGATGTTTTCAGGAGAAAGAAAGCTTAGGATATTTTGTGGTGT 585  
Qy 265 lAspMetAspAsnProIleGlyAsnTyrAspGlyArgPheAspGlyVal---LeuCysSe 284  
Db 586 GGACATGGATTAACCTATTGGCACTGGGATGGAGATTGATGGAGTGCAGCTTTGTAG 645  
Qy 284 rPheAlaCysValGluSerThrIleLeuLeuHisIleAsnAspIleIleProGluSerVa 304  
Db 646 TTTTGGCGTGTGTTGAAGTACAAATCTATTGCACATCAATGATATCATCCAGAGAGTGT 705  
Qy 304 lThrGlnGluArgProProLysLeuAlaPheMetSerArgGlyValGlyAspLysGI 324  
Db 706 GACGAGGAAAGGAGGCTCCCAACTTGCCTTTATGTCAAGAGGTGTGGGGACAAAGG 765  
Qy 324 ySerSerSerHisAsnLysProLysAlaThrGlySerThrSerAspProGlyAsnArg-- 343  
Db 766 TTCTATCCAGTCAATAAACCAGGCTACAGGATCTACCTCAGACCCTGGAAATAGAAA 825  
Qy 344 -ArgSerGluLeuPheTyrThrLeuAsnGlySerSerValAspSerGlnProGlnSerly 363  
Db 826 CAGATCTGAATTATTTTATACCTTAATGGGTCTTCTGTGACTCACAAACCAATCCAA 885  
Qy 363 sSerLysAsnThrTrpTyrIleAspGluValAlaGluAspProAlaLysSerLeuThrGI 383  
Db 886 ATCAAAAATACATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 945  
Qy 383 uileSerThrAspPheAspArgSerSerProLeuGlnProProValAsnSerIle 403  
Db 946 GATATCTACAGACTTTGACCGTCTTCACACCACTCCAGCCCTCTCTGTGAATCTACT 1005  
Qy 403 uThrThrGluAsnArgPheHisSerLeuProPheSerLeuThrLysMetProAsnThrAs 423  
Db 1006 GACCACCGAGAACAGATTCCACTCTTACCAATTCAGTCTCACCAAGATGCCCAATACCAA 1065  
Qy 423 nGlySerIleGlyHisSerProLeuSerLeuSerAlaGlnSerValMetGluGluLeuAs 443  
Db 1066 TGGAAATATTGGCCACAGTCCACTTCTCTGTGAGCCAGTCTGTAATGGAAAGAGCTAAA 1125  
Qy 443 nThrAlaProValGlnGluSerProProLeuAlaMetProProGlyAsnSerHisGlyIle 463  
Db 1126 CACTGCACCCGTCCAAGAGTCCACCTTGGCCATGCTCTGGGAATCTCATGGTCT 1185  
Qy 463 uGluValGlySerLeuAlaGluValLysGluAsnProProPheTyrGlyValIleArgTr 483  
Db 1186 AGAAGTGGCTCATTTGGCTGAAGTTAAGAGAACCTCTCTTCTATGGGTAATCCGTTG 1245  
Qy 483 pIleGlyGlnProProGlyLeuAsnGluValLeuAlaGlyLeuGluLeuGluAspGluCy 503  
Db 1246 GATCGGTGAGCCACAGGACTGAATGAAGTGTCTGTGAGCCAGTCTGGAATGGAAGAGTGT 1305  
Qy 503 sAlaGlyCysThrAspGlyThrPheArgGlyThrArgTyrPheThrCysAlaLeuLysIly 523  
Db 1306 TGCAGGCTGTACGAGTGAACCTTCAGAGGCACTCGGTATTTTCACTGTGCCCTGAAGAA 1365  
Qy 523 sAlaLeuPheValLysLeuLysSerCysArgProAspSerArgPheAlaSerLeuGlnPr 543  
Db 1366 GCGCGTGTGTGAAACCTGAAGAGCTGACGGCTGACTTAGTGTTCATCATTCGAGCC 1425  
Qy 543 oValSerAsnGlnIleGluArgCysAsnSerLeuAlaPheGlyGlyTyrLeuSerGluVa 563  
Db 1426 GGTTTTCCAATCAGATTGAGCGCTGTAACTCTTTAGCATTTGGAGGCTACTTAAGTGAAGT 1485

Qy 563 lValGluGluAsnThrProProLysMetGluLysGluGlyLeuGluIleMetIleGlyIy 583  
Db 1486 AGTAGAAGAAATAACTCCACCAAAATGGAAGAGAGGCTGGAGATAAATGATTGGGAA 1545  
Qy 583 slyLysGlyIleGlnGlyHisTyrAsnSerCysTyrLeuAspSerThrLeuPheCysIle 603  
Db 1546 GAAGAAAGGCAATCCAGGCTCATTAACAATTTCTTTACTTAGACTCAACCTTATCTGCTT 1605  
Qy 603 uPheAlaPheSerSerValLeuAspThrValLeu-LeuArgProLysGluLysAsnAspV 623  
Db 1606 ATTTGCTTTTAGTCTCTGTTCTGGACACTGTGTACTTGA-CCCAAGAAAGAAAGCATG 1664  
Qy 623 alGluTyrTyrSerGluThrGlnGluLeuLeuArgThrGluIleValAsnProLeuArgI 643  
Db 1665 TAGAATATTATAGTGAACCCCAAGAGCTACTGAGGACAGAAATTTGTTAATCTCTGAGAA 1724  
Qy 643 leTyrGlyTyrValCysAlaThrLysIleMetLysLeuArgLysIleLeuGluLysValG 663  
Db 1725 TATATGGGATATGTGTGTGCCCAAAATTTATGAAACTGAGGAAATATCTTGAAGAGTGG 1784  
Qy 663 luAlaAlaSerGlyPheThrSerGluGluLysAspProGluGluPheLeuAsnIleLeuP 683  
Db 1785 AGGCTGCATCAGATTACCTCTGAAGAAAGAAAGATCTTGAGGAATTTCTGAATATTTCTGT 1844  
Qy 683 heHisIleIleLeuArgValGluProLeuLeuLysIleArgSerAlaGlyGlnLysValG 703  
Db 1845 TTCAATCATATTTAAGGCTAGAACCTTTGCTTAAATAAAGATCAGCAGGTCAAAAGGTAC 1904  
Qy 703 lAspCysTyrPheTyrGlnIlePheMetGluLysAsnGluLysValGlyValProThrI 723  
Db 1905 AAGATTGTACTTCTATCAAAATTTTATGGAAGAAATATGAGAAAGTTGGCGTTCCACAA 1964  
Qy 723 leGlnGluLeuLeuTrpSerPheIleAsnSerAsnLeuLysPheAlaGluAlaProS 743  
Db 1965 TTCAGCAGTGTGTAGATGGTCTTTTATCAACAGTAACCTGAAATTTGCAGAGGACCAT 2024  
Qy 743 erCysLeuIleIleGlnMetProArgPheGlyLysAspPheLysLeuPheLysLysIleP 763  
Db 2025 CATGTCTGATTTATCAGATGCTCGATTTGGAAAGAGACTTTAAACTATTTTAAAAAATTT 2084  
Qy 763 heProSerLeuGluLeuAsnIleThrAspLeuLeuGluAspThrProArgGlnCysArgI 783  
Db 2085 TTCTCTCTCTGGAATTAATAATAACAGATTTTACTTGAAGACACTCCACAGACAGTCCGGA 2144  
Qy 783 leCysGlyGlyLeuAlaMetTyrGluCysArgGluCysTyrAspAspProAspIleSerA 803  
Db 2145 TATGTGGAGGGCTTCAATGTATGATGTAGAGATGTCTACCGATCCGGACATCTCAG 2204  
Qy 803 laGlyLysIleLysGlnPheCysLysThrCysAsnThrGlnValHisLeuHisProLysA 823  
Db 2205 CTGGAAGAAATCAAGCAGTTTGTAAACCTGCAACACTCAAGTCCACTTCAATCCGAGAA 2264  
Qy 823 tGLeuAsnHisLysTyrAsnProValSerLeuProLysAspLeuProAspTrpAspTrpA 843  
Db 2265 GCCTGAATCATAAATATAACCCAGTGTCACTTCCCAAGACTTACCCGACTGGGACTGGA 2324  
Qy 843 tGHisGlyCysIleProCysGlnAsnMetGluLeuPheAlaValLeuCysIleGluThrS 863  
Db 2325 GACAGGCTGCATCCCTTCCAGAAATATGAGTATTTTGTCTGTCTCTGCATAGAAACAA 2384  
Qy 863 erHisTyrValAlaPheValLysTyrGlyLysAspAspSerAlaTrpLeuPhePheAspS 883  
Db 2385 GCCACTATGTCTTTTGAAGATATGGGAAGACATTTCTGCCTGGCTCTTCTTTTGACA 2444  
Qy 883 erMetAlaAspArgAspGlyGlyGlnAsnGlyPheAsnIleProGlnValThrProCysP 903  
Db 2445 GCATGGCCGATCGGATGGTGTGATGGCTTCAACATTTCTCAAGTCAACCCCATGCC 2504  
Qy 903 roGluValGlyGluTyrLeuLysMetSerLeuGluAspLeuHisSerLeuAspSerArgA 923  
Db 2505 CAGAAAGTAGAGAGTACTTGAAGATGTCTCTGGAAGACCTGCATCTCTTGGACTCCAGGA 2564





Db 1306 TGCAGGCTGTACGGATGGAACTTCAGAGGCACTCGGTATTTCACCTGTGCGCCCTGAAGAA 1365  
Qy 523 sAlaLeuPheValLysLeuLysSerCysArgProAspSerArgPheAlaSerLeuGlnPr 543  
Db 1366 GCGCGTGTGTGAAACTGAAGAGCTGCAGGCTGACCTAGCTTAGCTTGCATCATTCGAGCC 1425  
Qy 543 oValSerAsnGlnIleGluArgCysAsnSerLeuAlaPheGlyGlyTyrLeuSerGluVa 563  
Db 1426 GGTTCCTCAATCAGATTGAGCGCTGTAACCTCTTTAGCAATTTGGAGGCTACTTAAAGTGAAGT 1485  
Qy 563 lValGluGluAsnThrProProLysMetGluLysGluGlyLeuGluIleMetIleGlyLys 583  
Db 1486 AGTAGAAGAAATACTCCACCAAAATGGAAAAGAGCTTTGGAGATAATGATTGGGAA 1545  
Qy 583 sLysLysGlyIleGlnGlyHisTyrAsnSerCysTyrLeuAspSerThrLeuPheCysLe 603  
Db 1546 GAAGAAAGCATCCAGGCTCATTAACATCTTGTACTTAGACTCAACCTTATCTGCTT 1605  
Qy 603 uPheAlaPheSerSerValLeuAspThrValLeu-LeuArgProLysGluLysAsnAspV 623  
Db 1606 ATTTGCTTTTATGTTCTGTCTTGACACACTGTGTACTTGA-CCCAAGAAAAGACGATG 1664  
Qy 623 alGluTyrTyrSerGluThrGlnGluLeuLeuArgThrGluIleValLeuProLeuArgI 643  
Db 1665 TAGAATATTATAGTGAACCCCAAGAGCTACTGAGGACAGAAATTGTTATCTCTGAGAA 1724  
Qy 643 leTyrGlyTyrValCysAlaThrLysIleMetLysLeuArgLysIleLeuGluLysValG 663  
Db 1725 TATATGGATATGTGTGGCCACAAAATTTATGAACTGAGGAAATATCTTTGAAAAGGTGG 1784  
Qy 663 luAlaAlaSerGlyPheThrSerGluGluLysAspProGluPheLeuAsnIleLeuP 683  
Db 1785 AGGCTGCATCAGATTACCTCTGAAGAAAAGATCCTGAGGAATCTTGAATATCTGT 1844  
Qy 683 heHisHisIleLeuArgValGluProLeuLeuLysIleArgSerAlaGlyGlnLysValG 703  
Db 1845 TTCATCATATTTTAAAGGTAGAACCTTTGCTTAAAAATTAAGATCAGCAGGTCAAAAGTAC 1904  
Qy 703 lnAspCysTyrPheTyrGlnIlePheMetGluLysAsnGluLysValGlyValProThrI 723  
Db 1905 AAGATTGTACTTCTATCAAAATTTTATGGAATAAATGAGAAAGTGTGGCGTTCCACAA 1964  
Qy 723 leGlnGlnLeuLeuGluTrpSerPheIleAsnSerAsnLeuLysPheAlaGluAlaProS 743  
Db 1965 TTCAGCAGTGTGTAGATGGTCTTTATCAACAGTAACTGAAATTTGCAGAGGCCACAT 2024  
Qy 743 erCysLeuIleIleGlnMetProArgPheGlyLysAspPheLysLeuPheLysIleP 763  
Db 2025 CATGTCTGATTATTCAGATGCCCTCGATTGGAAGAGACTTTAAACTATTTAAAAAAATTT 2084  
Qy 763 heProSerLeuGluLeuAsnIleThrAspLeuLeuGluAspThrProArgGlnCysArgI 783  
Db 2085 TTCTCTCTCTGAAATTAATAATAACAGATTATCTTGAAGACACTCCAGACAGTCCGGA 2144  
Qy 783 leCysGlyGlyLeuAlaMetTyrGluCysArgGluCysTyrAspAspProAspIleSerA 803  
Db 2145 TATGTGGAGGCTTCATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2204  
Qy 803 laGlyLysIleLysGlnPheCysLysThrCysAsnThrGlnValHisLeuHisProLysA 823  
Db 2205 CTGGAATAATCAAGCAGTTTGTAAAACTGCAACACTCAAGTCCACCTTCATCCGAGA 2264  
Qy 823 rgLeuAsnHisLysTyrAsnProValSerLeuProLysAspLeuProAspTrpAspTrpA 843  
Db 2265 GGCTGAATCATAAATATAACCAAGTGTCACTTCCCAAGACATTAACCGACTGGGACTGGA 2324  
Qy 843 rgHisGlyCysIleProCysGlnAsnMetGluLeuPheAlaValLeuCysIleGluThrS 863  
Db 2325 GACAGGCTGCATCTCTGCGCAAGATATGAGGATTTATGCTGCTTCTGTCATAGAAACAA 2384  
Qy 863 erHisTyrValAlaPheValLysTyrGlyLysAspSerAlaTrpLeuPheAspS 883  
Db 2385 GCCACTATGTTGCTTTTGTGAAGTATGGGAAGACGATTCTGCTGCTGCTTCTTTGACA 2444

Qy 883 erMetAlaAspArgAspGlyGlyGlnAsnGlyPheAsnIleProGlnValThrProCysP 903  
Db 2445 GCATGGCGATCGGATGGTGGTTCAGATGGCTTCAACATTCCTCAAGTACCCCATGCC 2504  
Qy 903 roGluValGlyGluTyrLeuLysMetSerLeuGluAspLeuHisSerLeuAspSerArgA 923  
Db 2505 CAGAAAGTAGGAGAGTACTTGAAGATGCTCTCGAAGAGCTGCAATCTCTGGAGTCCAGGA 2564  
Qy 923 rgIleGlnGlyCysAlaArgArgLeuLeuCysAspAlaTyrMetCysMetTyrGlnSerP 943  
Db 2565 GAATCCAAAGCTGTGCACGAAGACTGCTTTGTGATGCATATATGTCATGTACCAAGATC 2624  
Qy 943 roThrMetSerLeuTyrLys 949  
Db 2625 CAACAATGAGTTGTACAAA 2644

AK024348 2845 bp mRNA linear PRI 30-JAN-2004  
Homo sapiens cDNA FLJ14286 fis, clone PLACE1006129, highly similar  
to Homo sapiens HSPC057 mRNA.  
AK024348  
AK024348.1 GI:10436712  
oligo capping; fis (full insert sequence).  
Homo sapiens (human)  
SOURCE  
ORGANISM  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1  
Ota, T., Suzuki, Y., Nishikawa, T., Otsuki, T., Sugiyama, T., Irie, R.,  
Wakatsuki, A., Hayashi, K., Sato, H., Nagai, K., Kimura, K., Makita, H.,  
Sekine, M., Ohtsuka, M., Nishi, T., Shibahara, T., Tanaka, T.,  
Ishii, S., Yamamoto, J., Saito, K., Kawai, Y., Isono, Y., Nakamura, Y.,  
Nagahari, K., Murakami, K., Yasuda, T., Iwayanagi, T., Wagatsuma, M.,  
Shiratori, A., Sudo, H., Hosoiri, T., Kaku, Y., Kodaira, H., Kondo, H.,  
Sugawara, M., Takahashi, M., Kanda, K., Yokoi, T., Furuya, T.,  
Kikkawa, E., Omura, Y., Abe, K., Kamihara, K., Katsuta, N., Sato, K.,  
Tanikawa, M., Yamazaki, M., Ninomiya, K., Ishibashi, T., Yamashita, H.,  
Murakawa, M., Fujimori, K., Tanai, H., Kimata, M., Watanabe, M.,  
Hiraoka, S., Chiba, Y., Ishida, S., Oono, Y., Takiguchi, S., Watanabe, S.,  
Yosida, M., Hotuta, T., Kusano, J., Kanehori, K., Takahashi-Fujii, A.,  
Hara, H., Tanase, T., Nomura, Y., Togiyama, S., Komai, F., Hara, R.,  
Takeuchi, K., Arita, M., Imose, N., Musashino, K., Yuuki, H., Oshima, A.,  
Sasaki, N., Aotsuka, S., Yoshikawa, Y., Matsunawa, H., Ichihara, T.,  
Shohata, N., Sano, S., Moriya, S., Momiyama, H., Satoh, N., Takami, S.,  
Terashima, Y., Suzuki, O., Nakagawa, S., Senoh, A., Mizoguchi, H.,  
Goto, Y., Shimizu, F., Wakebe, H., Hishigaki, H., Watanabe, T.,  
Sugiyama, A., Takemoto, M., Kawakami, B., Yamazaki, M., Watanabe, K.,  
Kunagai, A., Itakura, S., Fukuzumi, Y., Fujimori, Y., Komiyama, M.,  
Tashiro, H., Tanigami, A., Fujiwara, T., Oho, T., Yamada, K., Fujii, Y.,  
Ozaki, K., Hirao, M., Ohmori, Y., Kawabata, A., Hikiji, T., Kobatake, N.,  
Inagaki, H., Ikema, Y., Okamoto, S., Okitani, R., Kawakami, T.,  
Noguchi, S., Itoh, T., Shigetani, K., Senba, T., Matsumura, K.,  
Nakajima, Y., Mizuno, T., Morinaga, M., Sasaki, M., Togae, T.,  
Oyama, M., Hata, H., Watanabe, M., Komatsu, T., Mizushima-Sugano, J.,  
Satoh, T., Shirai, Y., Takahashi, Y., Nakagawa, K., Okumura, K.,  
Nagase, T., Nomura, N., Kikuchi, H., Masuho, Y., Yamashita, R.,  
Nakai, K., Yada, T., Nakamura, Y., Ohara, O., Isogai, T., and Sugano, S.  
Complete sequencing and characterization of 21,243 full-length  
human cDNAs  
Nat. Genet. 36 (1), 40-45 (2004)  
14702039  
2  
Isogai, T., Ota, T., Hayashi, K., Sugiyama, T., Otsuki, T., Suzuki, Y.,  
Nishikawa, T., Nagai, K., Sugano, S., Shiratori, A., Sudo, H.,  
Wagatsuma, M., Hosoiri, T., Kaku, Y., Kodaira, H., Kondo, H.,  
Sugawara, M., Takahashi, M., Chiba, Y., Ishida, S., Murakawa, K.,  
Oono, Y., Takiguchi, S., Watanabe, S., Kimura, K., Murakami, K.,  
Ishii, S., Kawai, Y., Saito, K., Yamamoto, J., Wakatsuki, A.,  
Nakamura, Y., Nagahari, K., Masuho, Y., Ninomiya, K. and Iwayanagi, T.  
NEDO human cDNA sequencing project  
Unpublished



REFERENCE 3 (bases 1 to 2845)  
 AUTHORS Isogai,T. and Otsuki,T.  
 TITLE Direct Submission  
 JOURNAL Submitted (23-AUG-2000) Takao Isogai, Helix Research Institute,  
 Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan  
 (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)  
 COMMENT NEDO human cDNA sequencing project supported by Ministry of  
 International Trade and Industry of Japan; cDNA full insert  
 sequencing; Research Association for Biotechnology; cDNA library  
 construction; 5'- & 3'-end one pass sequencing and clone selection;  
 Helix Research Institute (supported by Japan Key Technology Center  
 etc.) and Department of Virology, Institute of Medical Science,  
 University of Tokyo.

## FEATURES

Location/Qualifiers  
 1..2845  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="PLACE1006129"  
 /tissue\_type="placenta"  
 /clone\_lib="PLACE1"  
 /note="Cloning vector: pME18SFL3"

## ORIGIN

## Alignment Scores:

Pred. No.: 0 Length: 2845  
 Score: 608.00 Matches: 861  
 Percent Similarity: 98.51% Conservative: 0  
 Best local Similarity: 98.51% Mismatches: 2  
 Query Match: 64.07% Indels: 13  
 DB: 9 Gaps: 0

US-09-671-687A-3 (1-949) x AK024348 (1-2845)

Qy 87 AspValValGluIleAsnGluIysPheThrGluLeuLeuAlaIleThrAsnCysGlu 106  
 Db 47 GATGTTGTAGATATAATGAAAGTTACAGAGTTACTTTTGGCAATTACCAATTGTGAG 106  
 Qy 107 GluArgPheSerLeuPheIysAsnArgAsnArgLeuSerIysGlyLeuGlnIleAspVal 126  
 Db 107 GAGAGGTTGAGGCTGTTTAAACACAGAAACAGACTAAGTAAGGAGGCTCCAAATAGACGTG 166  
 Qy 127 GlyCysProValIysValGlnLeuArgSerGlyGluLeuLysPheProGlyValValArg 146  
 Db 167 GCCTGCTCTGTGAAGTACAGCTGAGATCTGGGAAGAAAAATTTCTCGAGTTGACGC 226  
 Qy 147 PheArgGlyProLeuLeuAlaGluArgThrValSerGlyIlePhePheGlyValGluLeu 166  
 Db 227 TTCAGAGGACCCCTGTTAGCAGAGGACAGCTCCGGAAATATTTCTGGAGTTGAATTG 286  
 Qy 167 LeuGluGluGlyArgGlyGlnGlyPheThrAspGlyValTyrGlnGlyLysGlnLeuPhe 186  
 Db 287 CTGGAAGAAGGTCGTGGTCAAGGTTTCACTGACGGGGTGTACCAAGGAAACAGCTTTT 346  
 Qy 187 GlnCysAspGluAsnCysGly--PheValAlaLeuAspIysLeuIleLeuIleGluAsp 205  
 Db 347 CAGTGTGATGAAGATTGTGGCTGTTTGTGATTGGACAAGCTAGAACTCATAGAAGAT 406  
 Qy 206 AspAspThrAlaLeuGluSerAspTyrAlaGlyProGlyAspThrMetGlnValGluLeu 225  
 Db 407 GATGACACTGCATTGGAAAGTATTACGACGAGTCTTGGGGACACAAATGCGAGTTCGACTT 466  
 Qy 226 ProProLeuGluIleAsnSerArgValSerLeuLysGly-GlyGluThrIleGluSerGln 245  
 Db 467 CCTCTTTTGGAAATAAATCCACAGATTCTTTTGA--GGTTGGAGAAACAATAGAACTCTGG 525  
 Qy 245 YThrValIlePheCysAspValLeuProGlyLysGluSerLeuGlyTyrPheValGlyVa 265  
 Db 526 AACAGTTATATTCGTGTGATGTTTTCAGGAAAGAAAGCTAGGATATATTTTGTGGTGT 585  
 Qy 265 LAspMetAspAnProIleGlyAsnTrpAspGlyArgPheAspGlyVal---LeuCysSe 284  
 Db 586 GGACATGATTAACCTTATTGGCAACTGGGATGGAAGATTGTGATGGAGTGCAGCTTTGTAG 645

Qy 284 rPheAlaCysValGluSerThrIleLeuLeuHisIleAsnAspIleIleProGluSerVa 304  
 Db 646 TTTTGGCTGTGTTGAAAGTACAATTTCTATTGCACATCAATGATATCATCCAGAGAGTGT 705  
 Qy 304 lThrGlnGluArgArgProProIysLeuAlaPheMetSerArgGlyValGlyAspLysGln 324  
 Db 706 GAGCGAGGAAGAGAGGCTCCCAAACTTGCTTTTATGTCAAGAGGTGTGGGACAAAGG 765  
 Qy 324 YSerSerHisAsnLysProLysAlaThrGlySerThrSerAspProGlyAsnArg-- 343  
 Db 766 TTCATCCAGTCATAATAAACCAAGGCTACAGATCTACCTCAGACCTCGAAATAGAAA 825  
 Qy 344 -ArgSerGluLeuPheTyrThrLeuAsnGlySerSerValAspSerGlnProGlnSerLy 363  
 Db 826 CAGATCTGAATTATTTTATACCTTAAATGGGTCTTCTGTGACTCAACAACCAATCCAA 885  
 Qy 363 sSerLysAsnThrTyrTyrIleAspGluValAlaGluAspProAlaLysSerLeuThrGln 383  
 Db 886 ATCAAAAAATACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 945  
 Qy 383 uIleSerThrAspPheAspArgSerProProLeuGlnProProValAsnSerLe 403  
 Db 946 GATATCTACAGACTTTTGACCGTTCTTACCAACCACTCCAGCCTCTCTCTGTGAACCTCACT 1005  
 Qy 403 uThrThrGluAsnArgPheHisSerLeuProPheSerLeuThrLysMetProAsnThrAs 423  
 Db 1006 GACCACCGAGACAGATTTCATCTTTTACCATTCAGTCTCACCAGAGTGCCTCAATACCAA 1065  
 Qy 423 nGlySerIleGlyHisSerProLeuSerLeuSerAlaGlnSerValMetGluGluLeuAs 443  
 Db 1066 TGAAGTATTGGCCACAGTCCACTTTCTCTGTGAGCCAGTCTGTAATGGAAGAGCTAAA 1125  
 Qy 443 nThrAlaProValGlnGluSerProProLeuAlaMetProProGlyAsnSerHisGlyLe 463  
 Db 1126 CACTGCAACCGTCCAAAGAGAGTCCACCCTTGGCCATGCTCTCTGGGAACCTCAGATGCT 1185  
 Qy 463 uGluValGlySerLeuAlaGluValLysGluAsnProProPheTyrGlyValIleArgTr 483  
 Db 1186 AGAAGTGGGCTCATTTGGCTGAAGTTAAGGAGAAACCTCTCTTTCTATGGGGTAAATCCGTTG 1245  
 Qy 483 pIleGlyGlnProProGlyLeuAsnGluValLeuAlaGlyLeuGluLeuGluAspGluCy 503  
 Db 1246 GATCGTCAACCCACAGGAGTGAATGAAGTCTCGCTGGACTGGAACCTGGAAGATGATG 1305  
 Qy 503 sAlaGlyCysThrAspGlyThrPheArgGlyThrArgTyrPheThrCysAlaLeuLysLy 523  
 Db 1306 TGCAGGCTGTACGGATGGAACCTTTCAGAGGCACTCGGTATTTTCACCTGTGCTGCTGAAGA 1365  
 Qy 523 sAlaLeuPheValLysLeuLysSerCysArgProAspSerArgPheAlaSerLeuGlnPr 543  
 Db 1366 GCGCTGTTTGTGAAACTGAAGAGCTGCAGGCTGACTCTAGGTTTTCATCTATTCGACCC 1425  
 Qy 543 oValSerAsnGlnIleGluArgCysAsnSerLeuAlaPheGlyGlyTyrLeuSerGluVa 563  
 Db 1426 GGTTCCTCAATCAGATTGACGCTGTAACTCTTTTAGCATTTGGAGGCTACTTAAGTGAAGT 1485  
 Qy 563 lValGluGluAsnThrProLysMetGluLysGluGlyLeuGluIleMetIleGlyLy 583  
 Db 1486 AGTAGAAGAAATACTCCACCAAAATGGAAGAAAGGCTTCGAGATAATGATTGGAA 1545  
 Qy 583 sLysLysGlyIleGlnGlyHisTyrAsnSerCysTyrLeuAspSerThrLeuPheCysLe 603  
 Db 1546 GAAGAAAGGCACTCCAGGTCATTACAAATCTTGTACTTAGACTCAACCTTATTTCTGCTT 1605  
 Qy 603 uPheAlaPheSerSerValLeuAspThrValLeu-LeuArgProLysGluLysAsnAspV 623  
 Db 1606 ATTTGCTTTTAGTTCCTGTTCTGGACACTGTGTTACTTGA--CCCAAGAAAGAACGATG 1664  
 Qy 623 aGluTyrTyrSerGluThrGlnGluLeuLeuArgThrGluIleValAsnProLeuArgI 643  
 Db 1665 TAGAATATTATAGTGAACCAAGAGCTACTGAGGACAGAAATTTGTTAATCTCTCTGAGAA 1724

Qy	643	leTyrGlyTyrValCysalaThrLysIleMetLysleuargLysIleLeuGluLysValG	663
Db	1725	TATATGGATATGTGTGCCCAAAATTTATGAACTGAGGAAATATCTTGAAGAGGTGG	1784
Qy	663	luAlaAaSerGlyPheThrSerGluGluLysAspProGluGluPheLeuAanIleLeuP	683
Db	1785	AGGCTGCATCAGATTTACCTCTGAAGAAAAGATCCTGAGGAATCTTGAATATCTGT	1844
Qy	683	heHisIleLeuArgValGluProLeuLeuLysIleAargSerAlaGlyGlnLysValG	703
Db	1845	TTTCATCATATTTTAAGGTAGAACCTTTGCTAAATAAAGATCAGAGTCAAGAGTAC	1904
Qy	703	lnAspCysTyrPheTyrGlnIlePheMetGluLysAsnGluLysValGlyValProThrI	723
Db	1905	AAGATTGTACTTCTATCAAAATTTTATGGAATAAATGAGAAAGTTGGCGTTCCACAA	1964
Qy	723	leGlnGlnLeuLeuGluTrpSerPheIleAsnSerAsnLeuLysPheAlaGluAlaProS	743
Db	1965	TTTCAGCAGTGTGTAGAAATGGTCTTTTATCAACAGTAACCTGAAATTTGCAGAGGCACCAT	2024
Qy	743	erCysLeuIleIleGlnMetProAiqPheGlyLysAspPheLysLeuPheLysIleLeuP	763
Db	2025	CATGCTCATATTCAGATGCTCGATTTGGAAGAGCTTTAAACTATTTAAATAATTT	2084
Qy	763	heProSerLeuGlnLeuAanIleThrAspLeuLeuGluAspThrProArgGlnCysArgI	783
Db	2085	TTCTCTCTCGAAATTAATAATAACAGATTTACTTTGAAGACACTCCACAGAGTCCGGA	2144
Qy	783	leCysGlyGlyLeuAlaMetTyrGluCysArgGluCysTyrAspAspProAspIleSerA	803
Db	2145	TATGTGGAGGCGCTTCAATGATGATGATGATGATGATGATGATGATGATGATGATG	2204
Qy	803	laGlyLysIleLysGlnPheCysLysThrCysAsnThrGlnValHisLeuHisProLysA	823
Db	2205	CTGGAAATTCAGCAGCTTTTGTAAACCTGCAACACTCAAGTCCACCTTCATCCGAGA	2264
Qy	823	rgLeuAsnHisLysTyrAsnProValSerLeuProLysAspLeuProAspTrpAspTrpA	843
Db	2265	GGCTGAATCATATAATATAACCCAGTGTCACTTCCCAAGACATTACCGACTGGACTGA	2324
Qy	843	rgHisGlyCysIleProCysGlnAsnMetGluLeuPheAlaValLeuCysIleGluThrs	863
Db	2325	GACACGGCTGATCCCTTGCAGAAATATGAGTATTTGCTCTTCTGTCATAGAACAA	2384
Qy	863	erHisTyrValAlaPheValLysTyrGlyLysAspSerAlaTrpLeuPheAspS	883
Db	2385	GCCACTATGTGCTTTGTGAAGTATGGGAAGACGATTTCTGCCCTGCTCTTCTTGACA	2444
Qy	883	erMetAlaAspArgAspGlyGlyGlnAsnGlyPheAsnIleProGlnValThrProCysP	903
Db	2445	GCAATGCCGATCGGATGGTGTGATGATGGCTTCAACATTCCTCAAGTACCCCATGCC	2504
Qy	903	roGluValGlyGluTyrLeuLysMetSerLeuGluAspLeuHisSerLeuAspSerArgA	923
Db	2505	CAGAAGTAGGAGAGTACTTGAAGATGTCTCTGGAAGACCTGCATCTCTTGAGACTCCAG	2564
Qy	923	rgIleGlnGlyCysAlaArgArgLeuLeuCysAspAlaTyrMetCysMetTyrGlnSerP	943
Db	2565	GAATCAAGGCTGTCCAGAACACTGCTTTGTGATGCATATATGTCATGTACAGAGTC	2624
Qy	943	roThrMetSerLeuTyrLys	949
Db	2625	CAACATGATTTGTACAAA	2644
RESULT 12			
CO834226			
LOCUS			
DEFINITION			
SEQUENCE 97 from Patent WO2004058805.			
ACCESSION			
VERSION			
CO834226.1 GI:50833763			
KEYWORDS			
SOURCE			
Homo sapiens (human)			
ORGANISM			
Homo sapiens			

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
REFERENCE	1
AUTHORS	Mateuda, A. and Yoneta, S.
TITLE	T cell activating gene
JOURNAL	Patent: WO 2004058805-A 97 15-JUL-2004; Asahi Kasei Pharma Corporation (JP)
FEATURES	Location/Qualifiers
source	1..3311
	/organism="Homo sapiens"
	/mol_type="unassigned DNA"
	/db_xref="taxon:9606"
	243..3113
	/note="unnamed protein product"
	/codon_start=1
	/protein_id="CAH05328.1"
	/db_xref="GI:50833764"
	/translation="MSSGLWSEKVTSPYBERIPYLLQBCSVTDKQTQKLLKVPKG SIGQYIQRSVGHRSIPSAKGKQNGIKLQEPHVLFDKDYVEINEKFTELLA ITNCEERESLFKNRNLKGLQIDVGCVKQLRSGEKFGVVRFPGLAERTVSG IPFGVELLEGRGGFTDGVQKQFQDCDEDCGVFVADKLLELIEDDDTLAERTVAG PGDTMVELPLEINRSYSLKVGETIESGTVPFCDVLPKESLGYFVGVDMDNPIGNW DORFDVQLCSFACVESTILLINDIIPALSESVTQERRPPKLAIRMGVGDGKSSSH NKPKATGSTDPGNRRSELFTYLLNGSVDSQPSQSKNTWYIDEVADPAKSLTEIS TDFRSPPLQPPVNSLTENRHSLSFSLTKMENTNGSIGHLSLSAQSVMBELN TAPQESPLAMPNGNSHLEVLGSLAEVKNPPFYGVIRWIGQPLNEVLAGELED ECAGCTDGTFRGTFTCALKKALFKLSCRPDSRFSASLOPVSNQIRCNLSAFGGY LSEVVENTPPKMEKEGLEIMLKKGIQGHNSCYLSDTLFLCFAPFSVLDTLLRP KEKDVEYISETQELLRTIEVNPRLIYGVCAIKMKUKKILEKVEASGFTSESKDP BEFLNLFPHILRLVPLLIKIRSAQKQVQDCYFYQIFMEKNEKVGYPTTQQLLEWSPFN SNLPAEAPASCLIIQMPFRGDKFLKIPFSLNLNITDLELTDPRQRCICGLWAMY CRECYDDPDISAGKIQKFCCTKTOVHLHPKLNHKNYPSLKPDLPDMDWRHGCI PC QNMELFVLCIETSHYVAFVKYKDDSAWLFEDSMADRGONGFNIQVPTCPVEGE YLKWSLEDLHSLDSRRIOGCARRLLCDAYMCMYQSPMTSLYK"
ORIGIN	
Alignment Scores:	
Pred. No.:	0
Score:	598.00
Percent Similarity:	97.43%
Best Local Similarity:	97.43%
Query Match:	63.01%
DB:	6
	Length: 3311
	Matches: 947
	Conservative: 0
	Mismatches: 2
	Indels: 25
	Gaps: 0
US-09-671-687A-3 (1-949) x CO834226 (1-3311)	
Qy	1 MetSerSerGlyLeuTrpSerGlnGluLysValThrSerProTyrTrpGluGluArgile 20
Db	243 ATGAGTTCAGGCTTATGAGCCCAAGAAAAGTCACTTCACCCCTACTGGGAGAGCGGATT 302
Qy	21 PheTyrLeuLeuGlnGluCysSerValThrAspLysGlnThrGlnLysLeuLys 40
Db	303 TTTTACTTGTCTTCTTCAAGAAATGACGCGTTACAGACAAACAAACAAAGCTCTCTAAA 362
Qy	41 ValProLysGlySerIleGlnTyrIleGlnAspArgSerValGlyHisSerArgile 60
Db	363 GTACCAAGGGAAGTATAGGACAGTATATCAAGATCGTTCGTGGGCAATTCAGAGATT 422
Qy	61 ProSerAlaLysGlyLysLysAsnGlnIleGlyLeuLysIleLeuGluGlnProHisAla 80
Db	423 CTTCTGCAAAAGGCAAGAAAATTCAGATTGATTAATAATTTAGAGCAACTCATGCA 482
Qy	81 ValLeuPheValAspGlu---AspValValGluIleAsnGluLysPheThrGluLeuLeu 99
Db	483 GTTCTCTTTGTGTAAGAAAGGATGTTGTAGAGATAAATAAGTGAAGTTCACAGAGTTACTT 542
Qy	100 LeuAlaIleThrAsnCysGluGluArgPheSerLeuPheLysAsnArgAsnArgLeuSer 119
Db	543 TTGGCAATTAACAATTTGTGAGAGAGGTTTCAGCCTGTTTAAATAACAGAAACAGACTAAGT 602
Qy	120 LysGlyLeuGlnIleAspValGlyCysProValLysValGlnLeuArgSerGlyGluGlu 139
Db	603 AAAGGCTTCCAANTAGCTGGGCTGTCTGTGAAAGTACAGCTGAGATCTGGGGAAGAA 662

QY 140 LysPheProGlyValValArgPheArgGlyProLeuLeuAlaGluArgThrValSerGly 159  
DB 663 AAATTTCTGGAGTTGTACGCTTCAGAGGACCCCTGTAGCAGAGGACAGTCTCCGGA 722  
QY 160 IlePhePheGlyValGluLeuLeuGluGluGlyArgGlyGlnGlyPheThrAspGlyVal 179  
DB 723 ATATTCTTTGGAGTTGAATTTGCTGGAAGAAGGTCGTGGTCAAGGTTTCACTGACGGGGTG 782  
QY 180 TyrGlnGlyLysGlnLeuPheGlnCysAspGluAspCysGly---PheValAlaLeuAsp 198  
DB 783 TACCAGGGAACAGCTTTTCAGTGTGATGAAGATTGGCGGTGTGTGTCATTGGAC 842  
QY 199 LysLeuGluLeuIleGluAspAspThrAlaLeuGluSerAspTyrAlaGlyProGly 218  
DB 843 AAGCTAGAACTCATAGAGATGATGACACTGCTTGGAAAGTATTCAGCAGGTCCTGGG 902  
QY 219 AspThrMetGlnValGluLeuProLeuGluIleAsnSerArgValSerLeuLysGly 238  
DB 903 GACACAATGTCAGGTCGAACTTCCTCTCTTTGGAAATAAACTCCAGAGTTCCTTTGAA-GGT 961  
QY 239 -GlyGluThrIleGluSerGlyThrValIlePheCysAspValLeuProGlyLysGluSe 258  
DB 962 TGGAGAAACAATAGAACTCGGAACAGTTATTTCTGTGATGTTTGGCCAGGAAAGAAAG 1021  
QY 258 rLeuGlyTyrPheValGlyValAspMetAspAsnProIleGlyAsnTrpAspGlyArgPh 278  
DB 1022 CTAGAGATATTTGTGTGTGACATGATGATTAACCTATTGGCACTGGGATGGAAGATT 1081  
QY 278 eAspGlyVal---LeuCysSerPheAlaCysValGluSerThrIleLeuLeuHisIleAs 297  
DB 1082 TGATGGAGTGCAGCTTTGTAGTTTTCGGTGTGTGAAAGTACAAATTTCTATTCACATCAA 1141  
QY 297 nAspIleIlePro-----GluSerValThrGlnGluArgArgProProLysLeuAl 314  
DB 1142 TGATATCATCCAGCTTTTATCAGAGAGTGTGACGCAAGGAAAGAGGAGGCTCCCAAACTGC 1201  
QY 314 aPheMetSerArgGlyValGlyAspLysGlySerSerSerHisAsnLysProLysAlaTh 334  
DB 1202 CTTTATGTCAAGAGTGTGGGGACAAAGGTTTCATCCAGTCATAATAAACCAAGGCTAC 1261  
QY 334 rGlySerThrSerAspProGlyAsnArg---ArgSerGluLeuPheTyrThrLeuAsnG 353  
DB 1262 AGGATCTACTCAGACCTCGAAATAGAAACAGATCTGAATATTTTATACCTTAATGG 1321  
QY 353 ySerSerValAspSerGlnProGlnSerLysSerLysAsnThrTrpTyrIleAspGluVa 373  
DB 1322 GTCCTCTGTGACTCACCAACCAATCCAAATCAAAATAATACATGTTACATGATGAAGT 1381  
QY 373 lAlaGluAspProAlaLysSerLeuThrGluIleSerThrAspPheAspArgSerSerPr 393  
DB 1382 TGCAGAAGCCCTGCAAAATCTTTACAGAGATATCTACAGACTTTTGACCGTTCTTCACC 1441  
QY 393 oProLeuGlnProProValAsnSerLeuThrGluAsnAtqPheHisSerLeuPr 413  
DB 1442 ACCACTCCAGCCTCTCTCTGGAACTCACTGACCCAGAACAGATTCACCTCTTTACC 1501  
QY 413 oPheSerLeuThrLysMetProAsnThrAsnGlySerIleGlyHisSerProLeuSerLe 433  
DB 1502 ATTCACTCTACCAAGATGCCCAATACCAATGGAAGTATTGGCCACAGTCCACTTCTCT 1561  
QY 433 uSerAlaGlnSerValMetGluGluLeuAsnThrAlaProValGlnGluSerProProLe 453  
DB 1562 GTCAGCCCGAGTCTGAATGGAAAGAGCTAAACACTGCACCCGCTCCAAAGAGAGTCCACCTT 1621  
QY 453 uAlaMetProProGlyAsnSerHisGlyLeuGluValGlySerLeuAlaGluValLysG 473  
DB 1622 GGCCATGCTCTCTGGAACTCATGCTCTAGAAGTGGGCTCATTTGGCTGAAGTTAAGGA 1681  
QY 473 uAsnProProPheTyrGlyValIleArgTrpIleGlyGlnProProGlyLeuAsnGluVa 493  
DB 1682 GAACCCCTCTCTATGGGGTAAATCCGTTGGATCGGTGAGCCACCAAGGACTGAATGAAGT 1741

QY 493 lLeuAlaGlyLeuGluLeuLeuGluAspGluCysAlaGlyCysThrAspGlyThrPheArgG 513  
DB 1742 GCTCGCTGACTGGAACATGGAAGATGAGTGTGAGGCTGTACGATGGAACCTTCAGAGG 1801  
QY 513 yThrArgTyrPheThrCysAlaLeuLysLysAlaLeuPheValLysLeuLysSerCysAr 533  
DB 1802 CACTCGGTATTTCACTGTGCTGGAAGAGGCGCTGTTGTGAACTGAAGAGCTGCAG 1861  
QY 533 gProAspSerArgPheAlaSerLeuGlnProValSerAsnGlnIleGluArgCysAsnSe 553  
DB 1862 GCCTGACTCTAGGTTTGATCATTTGCGCGGTTTCCAATCAGATTGAGCGCTGTAATC 1921  
QY 553 rLeuAlaPheGlyGlyTyrLeuSerGluValValGluGluAsnThrProProLysMetG 573  
DB 1922 TTTAGCATTTGGAGGCTACTTAAGTGAAGTAGTAGAAGAAATACTCCACCAAAATGGA 1981  
QY 573 uLysGluGlyLeuGluIleMetIleGlyLysLysLysGlyIleGlnGlyHisTyrAnSe 593  
DB 1982 AAAAGAAGGCTTGGAGATAATGATTGGGAAGAAAGGCAATCCAGGCTCATTTACAATTC 2041  
QY 593 xCysTyrLeuAspSerThrLeuPheCysLeuPheAlaPheSerSerValLeuAspThrVa 613  
DB 2042 TTGTTACTTAGACTCAACCTTATTCGCTTATTTGCTTTTAGTCTGTTCTGGACACTGT 2101  
QY 613 lLeuLeuArgProProLysGluLysAsnAspValGluTyrTyrSerGluThrGlnGluLeu 633  
DB 2102 GTTACTTAGACCCAAAGAAAGAACGATGTAGAAATATTATAGTGAAACCCCAAGGCTACT 2161  
QY 633 uArgThrGluIleValAsnProLeuArgIleTyrGlyTyrValCysAlaThrLysIleMe 653  
DB 2162 GAGGACAGAAATGTTTAATCTCTGAGAAATATATGATATGTGTGTGCCACAAAAATTAT 2221  
QY 653 tLysLeuArg-LysIleLeuGluLysValGluAlaAlaSerGlyPheThrSerGluGluL 673  
DB 2222 GAAACTGAA-GAAAACTCTGAAAGGTGGAGCTGCATCAGATTTTACCTCTGAAGAAA 2280  
QY 673 yAspProGluGluPheLeuAsnIleLeuPheHisIleLeuArgValGluProLeuL 693  
DB 2281 AGATCTCAGGAAATTTCTGAATATTTCTGTTTCATCATATTTTAAGGTAGAACCTTTGC 2340  
QY 693 eUlysIleArgSerAlaGlyGlnLysValGlnAspCysTyrPheTyrGlnIlePheMetG 713  
DB 2341 TAAATAATAGATCAGCAGGTCAAAAGGTACAGATTGTACTTCTATCAATTTTATGG 2400  
QY 713 lLysAsnGluLysValGlyValProThrIleGlnGlnLeuLeuGluTrpSerPheIleA 733  
DB 2401 AAAAAAATGAGAAGTTGGCGTTCCCACAATTCAGCAGTTGTTAGAAATGCTTTTATCA 2460  
QY 733 snSerAsnLeuLysPheAlaGluAlaProSerCysLeuIleIleGlnMetProArgPheG 753  
DB 2461 ACAGTAACCTGAAATTTGAGAGGACCATCATGTCTGATTATTTCAGATGCCGATTTG 2520  
QY 753 lYlysAspPheLysLeuPheLysLysIlePheProSerLeuGluLeuAsnIleThrAspL 773  
DB 2521 GAAAGACTTTAAACTATTATAAAAAATTTTCTTCTGGAATTAATAATAACAGATT 2580  
QY 773 eUleuGluAspThrProArgGlnCysArgIleCysGlyGlyLeuAlaMetTyrGluCysA 793  
DB 2581 TACTTGAAGACACTCCACAGAGTGCAGGATATGTGAGGCTTGCAATGTATGATGATGA 2640  
QY 793 rgGluCysTyrAspAspProAspIleSerAlaGlyLysIleLysGlnPheCysLysThrC 813  
DB 2641 GAGAATGCTACGACGATCCGACATCTCAGCTGGAATAATCAAGCAGTTTGTAAACCT 2700  
QY 813 yAsnThrGlnValHisLeuHisProLysArgLeuAsnHisLysTyrAsnProValSerL 833  
DB 2701 GCACAACTCAGTCCACCTTCATCCGAGAGGCTGAATCATATAATATACCCAGTGTCA 2760  
QY 833 eProLysAspLeuProAspTrpAspTrpArgHisGlyCysIleProCysGlnAsnMetG 853  
DB 2761 TTCCCAAGACTTACCCGACTGGGACTGGAGACACGCTGCATCCCTTGCAGAAATATGG 2820  
QY 853 lLeuPheAlaValLeuCysIleGluThrSerHisTyrValAlaPheValLysTyrGlyL 873

```
Db 2821 AGTTATTGCTGTTCTCTGCTAGAAACAAGCCACTATGTTGCTTTTGTGAATGGA 2880
Qy 873 YSASPASPSeAlaTirPhePheAspSerMetAlaAspArgGlyGlyGlnAaNG 893
Db 2881 AGGACGATTCTGCTGCTGCTCTCTTTGACGATGCGCCGATGGGTGTCAGAAATG 2940
Qy 893 lyPheAenIleProGlnValThrProCysProGluValGlyGlyLysMetSerL 913
Db 2941 GCTTCAACATTCCTCAAGTCAACCCATGCCAGAGTAGGAGTACTTGAAGATGCTC 3000
Qy 913 euGluAapLeuHiSerLeuAspSerArgArgIleGlnGlyCysAlaArgLeuLeu 933
Db 3001 TGGAAGACCTGCATTCCTTGGACTCCAGGAAATCCAAAGGCTGTGCACGAGACTGCTTT 3060
Qy 933 YSASPAlaTyMetCysMetTyGlnSerProThrMetSerLeuTyLys 949
Db 3061 GTGATGCATATATGTCATGATACCCAGAGTCCAAACATGAGTTGTACAAA 3110

RESULT 13
LOCUS AK000187 2341 bp mRNA linear PRI 13-SEP-2003
DEFINITION Homo sapiens cDNA FLJ20180 fis, clone COL10238, highly similar to
AB020656 Homo sapiens mRNA for KIAA0849 protein.
ACCESSION AK000187
VERSION AK000187.1 GI:7020107
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Kawabata A., Hikiiji T., Kobatake N., Inagaki H., Ikema Y.,
Okamoto S., Okitani R., Ota T., Suzuki Y., Obayashi M., Nishi T.,
Shibahara T., Tanaka T., Nakamura Y., Isogai T. and Sugano S.
NEDO human cDNA sequencing project
Unpublished
2 (bases 1 to 2341)
Sugano S., Suzuki Y., Ota T., Obayashi M., Nishi T., Isogai T.,
Shibahara T., Tanaka T. and Nakamura Y.
Direct Submission
Submitted (15-FEB-2000) Sumio Sugano, Institute of Medical Science,
University of Tokyo, Department of Virology, Shirokane-dai, 4-6-1,
Minato-ku, Tokyo 108-8639, Japan (E-mail: flicdn@ims.u-tokyo.ac.jp,
Tel: 81-3-5449-5286, Fax: 81-3-5449-5416)
NEDO human cDNA sequencing project supported by Ministry of
International Trade and Industry of Japan; cDNA full insert
sequencing: Research Association for Biotechnology; cDNA library
construction, 5'- & 3'-end one pass sequencing: Department of
Virology and Human Genome Center, Institute of Medical Science,
University of Tokyo (partly supported by Science and Technology
Agency).
FEATURES             Location/Qualifiers
     source           1..2341
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="COL10238"
                     /tissue_type="colon"
                     /clone_lib="COL"
                     /note="Cloning vector pME18SFL3"
     misc_feature     1..2341
                     /note="highly similar to AB020656 Homo sapiens mRNA for
KIAA0849 protein"

ORIGIN
Alignment Scores:
Pred. No.:          0
Score:              517.00
Percent Similarity: 99.23%
Best Local Similarity: 99.23%
Query Match:       54.48%
DB:                 9
Length:             2341
Matches:            644
Conservative:       0
Mismatches:         1
Indels:              5
Gaps:               0
```

```
US-09-671-687A-3 (1-949) x AK000187 (1-2341)
Qy 305 ThrGlnGluArgArgProProLysLeuAlaPheMetSerArgGlyValGlyAspLysGly 324
Db 2 ACAGAGAAAGAGAGAGCCCTCCCAAACTTGCTTTATGTCAAGAGGTGTTGGGACAAAGT 61
Qy 325 SerSerSerHisAenLysProLysAlaThrGlySerThrSerAspProGlyAenArg--- 343
Db 62 TCATCCAGTCATAATAAACAAGAGCTACAGATCTACCTCAGACCTCGGAAATAGAAAC 121
Qy 344 ArgSerGluLeuPheTyThrLeuAenGlySerSerValAspSerGlnProGlnSerLys 363
Db 122 AGATCTGAATATATTTTATACCTTAAATGGGTCTTCTGTTGACTCAACAACCAATCCAAA 181
Qy 364 SerLysAenThrTrpTyIleAspGluValAlaGluAapProAlaLysSerLeuThrGlu 383
Db 182 TCAAAAAATATACATGATGTAAGTTGCAGAAAGACCCCTGCAAAATCTCTTACAGAG 241
Qy 384 IleSerThrAspPheAspArgSerSerProProLeuGlnProProValAenSerLeu 403
Db 242 ATATCTACAGACTTGGACCGTTCTTCACCACCACCTCCAGCCCTCCTCTGTGAACTCAGTG 301
Qy 404 ThrThrGluAenArgPheHisSerLeuProPheSerLeuThrLysMetProAenThrAen 423
Db 302 ACCACCGAGAACAGATTCCACTCTTTTACATTCAGTCTCACCAGATGCCCAATACCAAT 361
Qy 424 GlySerIleGlyHisSerProLeuSerSerAlaGlnSerValMetGluGluLeuAen 443
Db 362 GGAAGTATGGCCACAGTCCACTTCTCTGTGAGCCAGTCTGTATATGGAAGAGCTAAAC 421
Qy 444 ThrAlaProValGlnGluSerProProLeuAlaMetProProGlyAenSerHisGlyLeu 463
Db 422 ACTGACCCCGTCCAGAGAGTCCACCTTGGCCATGCTCTCTGGAACTCACATGCTCTA 481
Qy 464 GluValGlySerLeuAlaGluValLysGluAenProProPheTyGlyValIleArgTrp 483
Db 482 GAAGTGGGCTCATTTGGCTGAAGTTAAGGAGAACCCCTCTTCTATATGGGGTAAATCGTTGG 541
Qy 484 IleGlyGlnProProGlyLeuAenGluValLeuAlaGlyLeuGluLeuGluAapLys 503
Db 542 ATCGGTACGCCACCAAGGACTGAATGAAGTGTCTGCTGGACTGGAACTGGAAAGATGATGT 601
Qy 504 AlaGlyCysThrAspGlyThrPheArgGlyThrArgTyPheThrCysAlaLeuLysLys 523
Db 602 CAGGCTGTAGGATGGAAACCTTCAGAGGCACTCGGTATTTTACCTGTGCCCTCGAAGAG 661
Qy 524 AlaLeuPheValLysLeuLysSerCysArgProAspSerArgPheAlaSerLeuGlnPro 543
Db 662 GCGCTGTTTGTGAACCTGAGAGCTGCAGGCTGACTCTAGGTTTGCATCATTTGCAGCG 721
Qy 544 ValSerAenGlnIleGluArgCysAenSerLeuAlaPheGlyGlyTyLeuSerGluVal 563
Db 722 GTTTCCAATCAGATTGAGCGCTGTAACTCTTTAGCATTTGGAGGCTACTTAAGTGAAGTA 781
Qy 564 ValGluGluAenThrProProLysMetGluLysGluGlyLeuGluIleMetIleGlyLys 583
Db 782 GTAGAAGAAAAATACTCCCAAAAAATGAAAAAAGAGGCTTGGAGATAATATGATTGGGAAG 841
Qy 584 LysLysGlyIleGlnGlyHisTyAenSerCysTyThrLeuAapSerThrLeuPheCysLeu 603
Db 842 AAGAAGCATCCAGGGTCATTACAAATTTCTTTACTTAGACTCAACCTTATCTGCTTA 901
Qy 604 PheAlaPheSerSerValLeuAapThrValLeuLeuArgProLysGluLysAenAspVal 623
Db 902 TTTGCTTTTGTCTGTTCTGGACACTGTGTTACTTAGACCCCAAGAAAGAACGATGTA 961
Qy 624 GluTyThrSerGluThrGlnGluLeuLeuArgThrGluIleValAenProLeuArgIle 643
Db 962 GAATATTATAGTGAACCCCAAGAGCTACTGAGGACAGAAATTTGTTAAATCTCTCGAATA 1021
Qy 644 TyrGlyTyThrValCysAlaThrLysIleMetLysLeuArgGlyIleLeuGluLysValGlu 663
```

Db 1022 TATGATATGCTGTGCCACAAAAATATGAACCTCAGGAAAAATCTTGAAGGTGGAG 1081  
Qy 664 AlaLaserGlyPheThr-SerGluGluLysAspProGluGluPheLeuAsnIleuPhe 683  
Db 1082 GCTGATCAGGATTTGC-CTCTGAAGAAAAAGATCTGAGGAATCTTGAATATTTCTGT 1140  
Qy 683 eHisIleLeuArgValGluProLeuLeuLysIleArgSerAlaGlyGlnLysValGI 703  
Db 1141 TCATCATATTTAAGGGTAGAACCTTTGCTAAAATAAGATCAGCAGGTCAAAAGGTACA 1200  
Qy 703 nAspCysTyrPheTyrGlnIlePheMetGluLysAsnGluLysValGlyValProThrI 723  
Db 1201 AGATGTTTACTTCTATCAAAATTTTATGAAAAAATGAGAAAGTTGGCGTTCCCAAT 1260  
Qy 723 eGlnGlnLeuLeuGluTyrPheIleAsnSerAsnLeuLysPheAlaGluAlaProSe 743  
Db 1261 TCAGCAGTGTGTAGATGCTCTTTTATCAACAGTAACCTGAAATTTGCGAGGCGCAATC 1320  
Qy 743 rCysLeuIleLeuGlnMetProArgPheGlyLysAspPheLysLeuPheLysIlePhe 763  
Db 1321 ATGTCGTATTTAGATGCCTCGATTTGGAAGAAGCTTTAACTATTTAAAAAATTTT 1380  
Qy 763 eProSerLeuGluLeuAsnIleThrAspLeuLeuGluAspThrProArgGlnCysArgI 783  
Db 1381 TCCTTCTCTGGAATTAATAATACAGATTTTACTTGAAGACACTCCAGACAGTCCCGAT 1440  
Qy 783 eCysGlyLeuAlaMetTyrGluCysArgGluCysTyrAspAspProAspIleSerAl 803  
Db 1441 ATGTGGAGGCTGCAATGTATGATGATGAGTAGAATGCTACGACGATCCGACATCTCAG 1500  
Qy 803 aGlyLysIleLysGlnPheCysLysThrCysAsnThrGlnValHisLeuHisProLysAr 823  
Db 1501 TGGAAAAATCAAGCAGTTTGTAAACCTGCAACACTCAAGTCCACCTTCATCCGAAGAG 1560  
Qy 823 gLeuAsnHisLysTyrAsnProValSerLeuProLysAspLeuProAspTrpAspTrpAr 843  
Db 1561 GCTGAATCATAAATATAACCCAGTGTCACTTCCCAAGACTTACCGACTGGGACTGGAG 1620  
Qy 843 gHisGlyCysIleProCysGlnAsnMetGluLeuPheAlaValLeuCysIleGluThrSe 863  
Db 1621 ACAGGCTGCATCCCTTGCAGATATGAGGATTTTGTCTCTGTCATGATAAACAAG 1680  
Qy 863 rHisTyrValAlaPheValLysTyrGlyLysAspSerAlaTrpLeuPheAspSe 883  
Db 1681 CCATATGTTGCTTTTGTGAAGTATGGGAAGACGATTCGCTGCTCTCTCTTTGACAG 1740  
Qy 883 rMetAlaAspArgAspGlyGlyGlnAsnGlyPheAsnIleProGlnValThrProCysPr 903  
Db 1741 CATGCCGATCGGATGGTGTGTCAGATGGCTTCAACATTCCTCAAGTCAACCCATGCC 1800  
Qy 903 oGluValGlyGluTyrLeuLysMetSerLeuGluAspLeuHisSerLeuAspSerArgAr 923  
Db 1801 AGAAGTAGAGAGTACTTGAAGATGCTCTCGAAGACCTGCTATTCCTGACTCCAGGAG 1860  
Qy 923 gIleGlnGlyCysAlaArgArgLeuLeuCysAspAlaTyrMetCysMetTyrGlnSerPr 943  
Db 1861 AATCCAAGCTGTGCAAGAACTGCTTTGTGATGCATATATGTGATGTACAGAGTCC 1920  
Qy 943 oThrMetSerLeuTyrLys 949  
Db 1921 AACATGAGTTGTACAAA 1939

## RESULT 14

BD160470  
LOCUS BD160470 2569 bp DNA linear PAT 17-JAN-2003  
DEFINITION Primer for synthesizing full-length cDNA and use thereof.  
ACCESSION BD160470  
VERSION BD160470.1 GI:27866228  
KEYWORDS JP 2002191363-A/15313.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 2569)  
AUTHORS Ota,T., Isogai,T., Nishikawa,T., Hayaashi,K., Saito,K., Yamamoto,J.,  
Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.  
TITLE Primer for synthesizing full-length cDNA and use thereof  
JOURNAL Patent: JP 2002191363-A 15313 09-JUL-2002;  
COMMENT HELIX RESEARCH INSTITUTE  
OS Homo sapiens (human)  
PN JP 2002191363-A/15313  
PD 09-JUL-2002  
PF 28-JUL-2000 JP 2000280990  
PI TOSHIO OTA, TAKAO ISOGAI, TETSUO NISHIKAWA, KOJI HAYASHI, KAORU  
SAITO,  
PI JUNICHI YAMAMOTO, SHIZUKO ISHII, TOMOYASU SUGIYAMA, AI WAKAMATSU,  
PI KEIICHI NAGAI, TETSUJI OTSUKI  
PC  
C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/ PC  
10,  
PC C12P21/02, C12Q1/68//C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC  
Primer for synthesizing full-length cDNA and use thereof PH Key  
Location/Qualifiers  
FT CDS (403)..(1830).  
FEATURES  
source  
1..2569  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
ORIGIN  
Alignment Scores:  
Pred. No.: 0 Length: 2569  
Score: 418.00 Matches: 585  
Percent Similarity: 98.15% Conservatives: 0  
Best Local Similarity: 98.15% Mismatches: 1  
Query Match: 44.05% Indels: 11  
DB: Gaps: 0  
US-09-671-687A-3 (1-949) x BD160470 (1-2569)  
Qy 88 ValValGluIleAsnGluLysPheThrGluLeuLeuAlaIleThrAsnCysGluGlu 107  
Db 1 GTTGTAGAGATAAATGAAAAGTTTCACAGATTAATTTTGGCAATTACCAATTGTGAGGAG 60  
Qy 108 ArgPheSerLeuPheLysAsnArgAsnArgLeuSerLysGlyLeuGlnIleAspValGly 127  
Db 61 AGGTTCCAGCTGTTTAAACACAGAAACAGACTAAGTAAAGGCTCCAAAATAGACGTGGC 120  
Qy 128 CysProValLysValGlnLeuArgSerGlyGluGluLysPheProGlyValValArgPhe 147  
Db 121 TGTCTGTGAAGTACAGCTGAGATCTGGGAAGAAAAATTTCTTGGAGTTGTACGCTTC 180  
Qy 148 ArgGlyProLeuLeuAlaGluArgThrValSerGlyIlePhePheGlyValGluLeuLeu 167  
Db 181 AGAGGACCCCTGTTAGCAGAGGACAGTCTCCGGAATATTTCTTGGAGTTGAATGCTG 240  
Qy 168 GluGluGlyArgGlyGlnGlyPheThrAspGlyValTyrGlnGlyLysGlnLeuPheGln 187  
Db 241 GAAGAGAGTCTGTTCAAGGTTTCACTGACGGGGTGTACCAAGGAAACAGCTTTTTCAG 300  
Qy 188 CysAspGluAspCysGly---PheValAlaLeuAspLysLeuGluLeuIleGluAspAsp 206  
Db 301 TGTGATGAAGATTGTGGCGTGTGTTGTTGATTTGGACAGCTAGAACTCATAGAGATCAT 360  
Qy 207 AspThrAlaLeuGluSerAspTyrAlaGlyProGlyAspThrMetGlnValGluLeuPro 226  
Db 361 GACACTGCATTGGAAGTATTACGACAGTCTCTGGGACACAAATGCAGGTTCGAACCTTC 420  
Qy 227 ProLeuGluIleAsnSerArgValSerLeuLysGly- GlyGluThrIleGluSerGlyTh 246  
Db 421 CCTTTGGAATAAATCCAGAGTTTCTTTTCAA- GGTGGAGAAACAAATAGAAATCTGGAAC 479  
Qy 246 rValIlePheCysAspValLeuProGlyLysGluSerLeuGlyTyrPheValGlyValas 266  
Db 480 AGTTATATTCTGTGATGTTTTCAGGAAAGAAAGCTTAGGATATATTTTGTGTGTGGGA 539

```
Qy 266 pMetAspAsnProfileGlyAsnTrpAspGlyArgPheAspGlyVal----LeuCysSerPh 285
Db 540 CATGGATAACCTATTGGCAACTGGATGGAGATTTCATGGAGTGCAGCTTTGTAGTTT 599
Qy 285 eAlaCysValGluSerThrIleLeuLeuHisIleAsnAspIleIleProGluSerValTh 305
Db 600 TGCCTGTGTGTTGAAAGTACAAATCTATTGCACATCAATGATATCATCCAGAGAGTGTGAC 659
Qy 305 rGlnGluArgArgProProLysLeuAlaPheMetSerArgGlyValGlyAspLysGlySe 325
Db 660 GCAGAAAGGAGGCTCCCAACTTTTATGTCAAGAGGTGTGGGCAAAAGGTTT 719
Qy 325 rSerSerHisAsnLysProLysAlaThrGlySerThrSerAspProGlyAsnArg---Ar 344
Db 720 ATCCAGTCATATAAACCAGGCTACAGGATCTACCTCAGACCTCGGAAATAGAAACAG 779
Qy 344 gSerGluLeuPheThrLeuAsnGlySerSerValAspSerGlnProGlnSerLysSe 364
Db 780 ATCTGAATATTATTATACCTTAAATGGGTCTTCTGTGACTCACAACCAATCCAAATC 839
Qy 364 rLysAsnThrTrpTrpIleAspGluValAlaGluAspProAlaLysSerLeuThrGluIl 384
Db 840 AAAAAATACATGGTACATATGATGAAGTTGCAGAGACCTCGCAAAATCTCTTACAGAGAT 899
Qy 384 eSerThrAspPheAspArgSerSerProProLeuGlnProProValAsnSerLeuTh 404
Db 900 ATCTACAGACTTGGACCGTTCTTACCACCACTCCAGCTCTCTCTGTGAACTCACTGAC 959
Qy 404 rThrGluAsnArgPheHisSerLeuProPheSerLeuThrLysMetProAsnThrAsnGl 424
Db 960 CACCGAGAACAGATTCCACTCTTTACCATTTAGTCTCACCAGAGTGTGTAATGGAGAGCTAAACAC 1079
Qy 424 ySerIleGlyHisSerProLeuSerLeuSerAlaGlnSerValMetGluLeuLeuAsnTh 444
Db 1020 AAGTATTGGCCACAGTCCACTTCTCTGTGAGCCAGCTGTGTAATGGAGAGCTAAACAC 1079
Qy 444 rAlaProValGlnGluSerProProLeuAlaMetProProGlyAsnSerHisGlyLeuGl 464
Db 1080 TGCACCGCTCCAGAGAGTCCACCTTGGCCATGCTCTGGGAATCTACATGGTCTAGA 1139
Qy 464 uValGlySerLeuAlaGluValLysGluAsnProProPheThrGlyValIleArgTrpIl 484
Db 1140 AGTGGCTCATTTGGCTGAAGTTAAGGAGAACCTCTCTTCTATGGGTAAATCCGTTGGAT 1199
Qy 484 eGlyGlnProProGlyLeuAsnGluValLeuAlaGlyLeuGluLeuGluAspGluCysAl 504
Db 1200 CGGTGAGCCAGCAGACTGAATGAAGTGTCTGCTGAGCTGGAACTGGAGATGAGTGTGC 1259
Qy 504 aGlyCysThrAspGlyThrPheArgGlyThrArgTyrrPheThrCysAlaLeuLysVal 524
Db 1260 AGCTGTACGATGAGACTTTCAGAGGACCTCGGTATTTCCTGTGCTGCTGAGAGGC 1319
Qy 524 aLeuPheValLysLeuLysSerCysArgProAspSerArgPheAlaSerLeuGlnProVa 544
Db 1320 GCTGTTGTGAACTGAAGAGCTGCAGGCTGACTCTAGGTTTGCATCATTCGACCGGT 1379
Qy 544 lSerAsnGlnIleGluArgCysAsnSerLeuAlaPheGlyGlyTyrrLeuSerGluValVa 564
Db 1380 TTCCAATCAGATTGAGCGCTGTAACTTTTAGCATTTGAGGCTACTTAAGTGAAGTAGT 1439
Qy 564 lGluGluAsnThrProProLysMetGluLysGluGlyLeuGluIleMetIleGlyLysIly 584
Db 1440 AGAAGAAATAACTCACCAAAATCGAAAAAGAGGCTTGGAGATAATGATTGGGAAGAA 1499
Qy 584 sLysGlyIleGlnGlyHisTyrrAsnSerCysTyrrLeuAspSerThrLeuPheCysLeuPh 604
Db 1500 GAAAGGCATCAGGGTCAATTAATCTTGTACTTAGACTCAACTTATTCTGCTTATT 1559
Qy 604 eAlaPheSerSerValLeuAspThrValLeuLeuArgProLysGluLysAsnAspValGl 624
Db 1560 TGCITTTAGTCTGTCTGGACACTGTGTACTTAGACCCCAAGAAAGAACGATGTAGA 1619
```

```
Qy 624 uTyrrTyrrSerGluThrGlnGluLeuLeuArgThrGluIleValAsnProLeuArgIleTy 644
Db 1620 ATATTATAGTGAACCCAGAGCTACTCAGGACAGAAATTTGTTAATCTCTGAGAATAATA 1679
Qy 644 rGlyTyrrValCysAlaThrLysIleMetLysLeuArgLysIleLeuGluLysValGluAl 664
Db 1680 TGGATATGTGTGTGCCACAAAATTTATGAAACTGAGGAAAACTTGAAGAGGTGGAGGC 1739
Qy 664 aAlaSerGlyPheThrSerGluGluLys 673
Db 1740 TGCATCAGATTACCTCTGAAGAAAAA 1767

RESULT 15
AX883681
LOCUS AX883681
DEFINITION Sequence 18596 from Patent EP1074617.
ACCESSION AX883681
VERSION AX883681.1 GI:40038582
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Ota, T., Isogai, T., Nishikawa, T., Hayashi, K., Saito, K., Yamamoto, J.,
Ishii, S., Sugiyama, T., Wakamatsu, A., Nagai, K. and Otsuki, T.
TITLE Primers for synthesising full-length cDNA and their use
JOURNAL Patent: EP 1074617-A 18586 07-FEB-2001;
Research Association for Biotechnology (JP)
FEATURES
source
1..2569
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="caxon:9606"
403..1833
/feature="unnamed protein product"
/codon_start=1
/protein_id="CAE92065.1"
/db_xref="GI:40038583"
/translation="MQVELPPELEINSRVLKVGETIESGTVIPCDVLPKESLGLFVGV
YDMDNPIGNWDKRGFDVGLQSCFACVETILLHINDIIPESVTQERRPPKLAIPMSRGVG
DKGSSHNKPKATGSDTPGNRNSRSELYTLNGSDVQSQPSKXNTWYIDVADPA
KSLTEITDTPDRSPPLQPPVNSLTENRFLSLPFLTKMPTNGSIGHSLSLAQ
SVMEELANTAPQESPPPLAMPNGSHGLEVSEIAEVKENPPFVGVIRWICQPPGLENV
AGLEDEBACACTDGTGRTFYTCALFKALFKVLSKCRPSRPFASLPVSNQIRCN
SLAFGYLUSEVEENTPPMBKGLIEMIKKKGIQGHNSCYLDSTLFCULFAPFSSVL
DTLLRPKKNDEYYSIQELLRLTEIVNPLRIYGVVCACTKIMKLRLKLEKVEAASGF
TSEBKGDHLNLYALKITBEIHMSKY"
ORIGIN
Alignment Scores:
Pred. No.: 0 Length: 2569
Score: 418.00 Matches: 585
Percent Similarity: 98.15% Conservative: 0
Best Local Similarity: 98.15% Mismatches: 1
Query Match: 44.05% Indels: 11
Db: 6 Gaps: 0

US-09-671-687A-3 (1-949) x AX883681 (1-2569)
Qy 88 ValValGluIleAsnGluLysPheThrGluLeuLeuAlaIleThrAsnCysGluGlu 107
Db 1 GTTGTAGAGATAAATGAAAAGTTTCACAGATTACTTTTGGCAATTTACCAATTTGTGAGGAG 60
Qy 108 ArgPheSerLeuPheLysAsnArgAsnArgLeuSerLysGlyLeuGlnIleAspValGly 127
Db 61 AGGTTACGCTGTTTAAAAACAGAAACAGACTAAGTAAGGCTCCAAATAGACGTGGGC 120
Qy 128 CysProValLysValGlnLeuArgSerGlyGluGluLysPheProGlyValValArgPhe 147
Db 121 TGTCTGTGAAAGTACAGCTGAGACTCTGGGAGAAAAATTTCTCGGAGTTGTACGCTTC 180
Qy 148 ArgGlyProLeuAlaGluArgThrValSerGlyIlePhePheGlyValGluLeuLeu 167
```



Db 181 AGAGACCCCTGTTAGCAGAGGACAGCTCCGGAAATATTTCTTGGAGTTGAATTCGCTG 240  
Qy 168 GluGluGlyArgGlyGlnGlyPheThrAspGlyValTyrGlnGlyLysGlnLeuPheGln 187  
Db 241 GAAGAAGGTGCGGGTCAAGGTTTCACTGACGGGGTGACCAAGGGAACAGCTTTTTCAG 300  
Qy 188 CysAspGluAspCysGly---PheValAlaLeuAspLysLeuGluLeuLeuGluAspAsp 206  
Db 301 TGTGATGAAGATTGTGGCGTGTGTTGTGATTGGACAAGCTAGAACTCATAGAGATGAT 360  
Qy 207 AspThrAlaLeuGluSerAspTyrAlaGlyProGlyAspThrMetGlnValGluLeuPro 226  
Db 361 GACACTGCATTGGAAAGTATTACGACGAGTCTCTGGGGACACAAATGCAGGTCTGAACTTCT 420  
Qy 227 ProLeuGluLeuAsnSerArgValSerLeuLysGly-GlyGluThrIleGluSerGlyTh 246  
Db 421 CCTTTGGAAATAAATCCACAGATTTCTTTGAA--GGTTGGAGAAACAATAGAAATCTGGAAC 479  
Qy 246 rValIlePheCysAspValLeuProGlyLysGluSerLeuGlyTyrPheValGlyValAs 266  
Db 480 AGTTATATTCGTGATGTTTCCAGGAAAGAAAGCTTAGGATATTTTGTGGTGTGA 539  
Qy 266 pMetAspAsnProIleGlyAsnTrpAspGlyArgPheAspGlyVal---LeuCysSerPh 285  
Db 540 CATGATAAACCTATTGGCAACTGGGATGGAAAGATTGTGAGTGCAGCTTTGTAGTTT 599  
Qy 285 eAlaCysValGluSerThrIleLeuLeuHisIleAsnAspIleIleProGluSerValTh 305  
Db 600 TCGGTGTGTGAAGTACAAATTCATTGTGCACATCATATGATATCATCCAGAGAGTGTGAC 659  
Qy 305 rGlnGluArgArgProProLysLeuAlaPheMetSerArgGlyValGlyAspLysGlySe 325  
Db 660 GAGGAAGAGGAGGCTCCCAAACTTGCCTTTATGTCAGAGGTTGGGGACAAAGGTTT 719  
Qy 325 rSerSerHisAsnLysProLysAlaThrGlySerThrSerAspProGlyAsnArg---Ar 344  
Db 720 ATCCAGTCATAATAACCAAGGCTACAGAGTCTACCTCAGACCTTGGAATAGAAACAG 779  
Qy 344 gSerGluLeuPheThrLeuAsnGlySerSerValAspSerGlnProGlnSerLysSe 364  
Db 780 ATCTGAATATTATTTATACCTTAATAGGTTCTCTGTGACTCACAAACCAACATCCAAATC 839  
Qy 364 rLysAsnThrTrpTyrIleAspGluValAlaGluAspProAlaLysSerLeuThrGluIl 384  
Db 840 AAAAAATACATGGTACATTGATGAAGTTGCAGAAGACCTCTGCAAAATCTCTTACAGAGAT 899  
Qy 384 eSerThrAspPheAspArgSerSerProProLeuGlnProProProValAsnSerLeuTh 404  
Db 900 ATCTACAGACTTTGACCGTCTTTCACCACTCCAGCTCTCTCTGTGAACCTCACTGAC 959  
Qy 404 rThrGluAsnArgPheHisSerLeuProPheSerLeuThrLysMetProAsnThrAsnGl 424  
Db 960 CACCGAGAACAGATTCCACTCTTTTACCATTAGTCTCACCAAGATGCCCAATACCAATGS 1019  
Qy 424 ySerIleGlyHisSerProLeuSerLeuSerAlaGlnSerValMetGluGluLeuAsnTh 444  
Db 1020 AAGTATTGGCCACAGTCCACTTCTCTGTGAGCCAGTCTGTGAATGGAAGAGCTAAACAC 1079  
Qy 444 rAlaProValGlnGluSerProProLeuAlaMetProProGlyAsnSerHisGlyLeuGl 464  
Db 1080 TGCACCCCTGCAAGAGTCCACCTTTGGCCATGCTCTCTGGAACTCACATGGTCTAGA 1139  
Qy 464 uValGlySerLeuAlaGluValLysGluAsnProProPheTyrGlyValIleArgTrpIl 484  
Db 1140 AGTGGCTCAITGGCTGAAGTTAAGAGAACCTCTCTTCTATGGGGTAATCCGTTGGAT 1199  
Qy 484 eGlyGlnProProGlyLeuAsnGluValLeuAlaGlyLeuGluLeuGluAspGluCysAl 504  
Db 1200 CGGTGAGCCACCGAGACTGAATGAAGTGTCTGCTGGACTTGGAACTGGAAGATGAGTGTGC 1259  
Qy 504 aGlyCysThrAspGlyThrPheArgGlyThrArgTyrPheThrCysAlaLeuLysLysAl 524

Db 1260 AGCCTGTACGGATGGAAACCTTTCAGAGGCACCTCGGTATTTTCCCTGTGCCCTGAAAGAGGC 1319  
Qy 524 aLeuPheValLysLeuLysSerCysArgProAspSerArgPheAlaSerLeuGlnProVa 544  
Db 1320 GCTGTTTGGAAACTGGAAGAGCTGCAGGCTGACTCTAGGTTTGCATCATTGAGCCGGT 1379  
Qy 544 lSerAsnGlnIleGluArgCysAsnSerLeuAlaPheGlyGlyTyrLeuSerGluValVa 564  
Db 1380 TTCCAATCAGATTGAGCGCTGTAACCTTTTAGCATTTGGAGGCTACTTTAAGTGAAGTAGT 1439  
Qy 564 lGluGluAsnThrProProLysMetGluLysGluGlyLeuGluIleMetileGlyLysLy 584  
Db 1440 AGAAGAAAAATACTCCACCAAAAATGGAAGAAAGAGGCTTGGAGATAATGATTGGGAAGAA 1499  
Qy 584 sLysGlyIleGlnGlyHisTyrAsnSerCysTyrLeuAspSerThrLeuPheCysLeuPh 604  
Db 1500 GAAAGGCATCCAGGGCTATTACAATTTCTGTTACTTAGACTCAACCTTATTCGCTTATT 1559  
Qy 604 eAlaPheSerSerValLeuAspThrValLeuLeuArgProLysGluLysAsnAspValGl 624  
Db 1560 TGCTTTTAGTTCTGTTCTGGACACTGTGTTACTTAGACCCCAAGAAAGAACGATGTAGA 1619  
Qy 624 uTyrTyrSerGluThrGlnGluLeuLeuArgThrGluIleValAsnProLeuArgIleTy 644  
Db 1620 ATATTATAGTGAACCCCAAGAGCTACTGAGGACAGAAATTTGTAATCTCTGAGAATATA 1679  
Qy 644 rGlyTyrValCysAlaThrLysIleMetLysLeuArgLysIleLeuGluLysValGluAl 664  
Db 1680 TGGATATGTGTGCCCAAAAATTTATGAACTTGAGGAAAAATACTTGGAAAAAGGTGGAGGC 1739  
Qy 664 aAlaSerGlyPheThrSerGluLys 673  
Db 1740 TGCATCAGGATTTTACCTCTGGAAGAAAA 1767

Search completed: April 17, 2005, 09:55:12

Job time : 9084 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 18, 2005, 09:37:54 ; Search time 355 Seconds  
(without alignments)  
3122.360 Million cell updates/sec

Title: US-09-671-687A-3  
Perfect score: 949  
Sequence: 1 MSSGLWSQEKVTPYWEERI.....RLLCDAYMCMYQSPMTSLYK 949

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 6959266 seqs, 116806243 residues

Word size : 0

Total number of hits satisfying chosen parameters: 6959266

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: listing first 45 summaries

Database : Pending Patents AA Main:\*

- 1: /cgn2\_6/ptodata/1/paa/pctus COMB.pcp.\*
- 2: /cgn2\_6/ptodata/1/paa/US06 COMB.pcp.\*
- 3: /cgn2\_6/ptodata/1/paa/US07 COMB.pcp.\*
- 4: /cgn2\_6/ptodata/1/paa/US08 COMB.pcp.\*
- 5: /cgn2\_6/ptodata/1/paa/US081 COMB.pcp.\*
- 6: /cgn2\_6/ptodata/1/paa/US082 COMB.pcp.\*
- 7: /cgn2\_6/ptodata/1/paa/US083 COMB.pcp.\*
- 8: /cgn2\_6/ptodata/1/paa/US084 COMB.pcp.\*
- 9: /cgn2\_6/ptodata/1/paa/US085 COMB.pcp.\*
- 10: /cgn2\_6/ptodata/1/paa/US086 COMB.pcp.\*
- 11: /cgn2\_6/ptodata/1/paa/US087 COMB.pcp.\*
- 12: /cgn2\_6/ptodata/1/paa/US088 COMB.pcp.\*
- 13: /cgn2\_6/ptodata/1/paa/US089 COMB.pcp.\*
- 14: /cgn2\_6/ptodata/1/paa/US090 COMB.pcp.\*
- 15: /cgn2\_6/ptodata/1/paa/US091 COMB.pcp.\*
- 16: /cgn2\_6/ptodata/1/paa/US092 COMB.pcp.\*
- 17: /cgn2\_6/ptodata/1/paa/US093 COMB.pcp.\*
- 18: /cgn2\_6/ptodata/1/paa/US094 COMB.pcp.\*
- 19: /cgn2\_6/ptodata/1/paa/US095 COMB.pcp.\*
- 20: /cgn2\_6/ptodata/1/paa/US096 COMB.pcp.\*
- 21: /cgn2\_6/ptodata/1/paa/US097A COMB.pcp.\*
- 22: /cgn2\_6/ptodata/1/paa/US097B COMB.pcp.\*
- 23: /cgn2\_6/ptodata/1/paa/US098 COMB.pcp.\*
- 24: /cgn2\_6/ptodata/1/paa/US099A COMB.pcp.\*
- 25: /cgn2\_6/ptodata/1/paa/US099B COMB.pcp.\*
- 26: /cgn2\_6/ptodata/1/paa/US100 COMB.pcp.\*
- 27: /cgn2\_6/ptodata/1/paa/US101 COMB.pcp.\*
- 28: /cgn2\_6/ptodata/1/paa/US102 COMB.pcp.\*
- 29: /cgn2\_6/ptodata/1/paa/US103 COMB.pcp.\*
- 30: /cgn2\_6/ptodata/1/paa/US104 COMB.pcp.\*
- 31: /cgn2\_6/ptodata/1/paa/US105 COMB.pcp.\*
- 32: /cgn2\_6/ptodata/1/paa/US106 COMB.pcp.\*
- 33: /cgn2\_6/ptodata/1/paa/US107 COMB.pcp.\*
- 34: /cgn2\_6/ptodata/1/paa/US108 COMB.pcp.\*
- 35: /cgn2\_6/ptodata/1/paa/US109 COMB.pcp.\*
- 36: /cgn2\_6/ptodata/1/paa/US110 COMB.pcp.\*
- 37: /cgn2\_6/ptodata/1/paa/US60 COMB.pcp.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	949	100.0	949	20	US-09-671-687A-3	Sequence 3, Appli
2	606	63.9	685	18	US-09-488-725A-2399	Sequence 2399, Ap
3	606	63.9	685	28	US-10-258-898A-2399	Sequence 2399, Ap
4	606	63.9	685	28	US-10-286-897-2399	Sequence 2399, Ap
5	606	63.9	731	22	US-09-786-797B-9	Sequence 9, Appli
6	606	63.9	731	35	US-10-921-707-9	Sequence 9, Appli
7	606	63.9	731	37	US-60-131-321-7	Sequence 7, Appli
8	606	63.9	739	1	PCT-US01-01239-1743	Sequence 1743, Ap
9	606	63.9	739	22	US-09-764-902-1743	Sequence 1743, Ap
10	606	63.9	953	1	PCT-US02-27777A-137	Sequence 137, App
11	606	63.9	953	1	PCT-US02-27777A-137	Sequence 137, App
12	606	63.9	953	27	US-10-170-205E-37570	Sequence 37570, A
13	606	63.9	956	1	PCT-US02-14570-4	Sequence 4, Appli
14	606	63.9	956	23	US-09-851-673-4	Sequence 4, Appli
15	606	63.9	956	33	US-10-755-889-490	Sequence 490, App
16	606	63.9	956	37	US-60-440-068-490	Sequence 490, App
17	606	63.9	956	37	US-60-469-757-490	Sequence 490, App
18	606	63.9	960	1	PCT-US04-07268-250	Sequence 250, App
19	606	63.9	960	33	US-10-788-792-250	Sequence 250, App
20	505	53.2	731	20	US-09-629-469A-18843	Sequence 18843, A
21	505	53.2	731	35	US-10-917-503-18843	Sequence 18843, A
22	363	38.3	698	18	US-09-488-725A-5971	Sequence 5971, Ap
23	363	38.3	698	28	US-10-258-898A-5971	Sequence 5971, Ap
24	363	38.3	698	28	US-10-286-897-5971	Sequence 5971, Ap
25	330	34.8	476	20	US-09-629-469A-18587	Sequence 18587, A
26	330	34.8	476	35	US-10-917-503-18587	Sequence 18587, A
27	227	23.9	558	22	PCT-US01-01239-1207	Sequence 1207, Ap
28	227	23.9	558	22	US-09-764-902-1207	Sequence 1207, Ap
29	208	21.9	261	1	PCT-US01-16450-1609	Sequence 1609, Ap
30	208	21.9	261	1	PCT-US01-16450A-1609	Sequence 1609, Ap
31	208	21.9	261	28	US-10-264-237-1609	Sequence 1609, Ap
32	192	20.2	394	1	PCT-US01-14827-12399	Sequence 12399, A
33	148	15.6	238	1	PCT-US01-16450-1610	Sequence 1610, Ap
34	148	15.6	238	1	PCT-US01-16450A-1610	Sequence 1610, Ap
35	148	15.6	238	28	US-10-264-237-1610	Sequence 1610, Ap
36	140	14.8	512	1	PCT-US01-14827-12400	Sequence 12400, A
37	118	12.4	145	1	PCT-US01-14827-12395	Sequence 12395, A
38	118	12.4	145	1	PCT-US01-14827-15308	Sequence 15308, A
39	113	11.9	113	1	PCT-US01-00663-27668	Sequence 27668, A
40	113	11.9	113	23	US-09-864-761-34675	Sequence 34675, A
41	113	11.9	113	27	US-10-182-993-28809	Sequence 28809, A
42	113	11.9	113	27	US-10-182-995-21147	Sequence 21147, A
43	113	11.9	113	27	US-10-182-997-13791	Sequence 13791, A
44	113	11.9	113	27	US-10-182-998-11431	Sequence 11431, A
45	113	11.9	113	28	US-10-203-134-27410	Sequence 27410, A

ALIGNMENTS

RESULT 1  
US-09-671-687A-3  
; Sequence 3, Application US/09671687A  
; GENERAL INFORMATION:  
; APPLICANT: WALLACH, David  
; APPLICANT: KOVALENKO, Andrei  
; APPLICANT: CANTARELLA, Giuseppina  
; TITLE OF INVENTION: INHIBITOR OF NF-KB ACTIVATOR  
; FILE REFERENCE: WALLACH=25  
; CURRENT APPLICATION NUMBER: US/09/671,687A  
; CURRENT FILING DATE: 2000-09-28  
; PRIOR APPLICATION NUMBER: 09/646,403  
; PRIOR FILING DATE: 2000-09-18  
; PRIOR APPLICATION NUMBER: IL 126024  
; PRIOR FILING DATE: 1998-09-01  
; PRIOR APPLICATION NUMBER: IL 134604  
; PRIOR FILING DATE: 2000-02-17  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: PatentIn version 3.1

```
; SEQ ID NO 3
; LENGTH: 949
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-671-687A-3

Query Match      100.0%; Score 949; DB 20; Length 949;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 949; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSGLWSQEKVTSPTWEERIFVLLLOECSVTDKQTKLLKVPKSGIGQYIDRSVGHRI 60
Db 1 MSSGLWSQEKVTSPTWEERIFVLLLOECSVTDKQTKLLKVPKSGIGQYIDRSVGHRI 60
QY 61 PSAKGKKNQIGIKILEQPHAVLFVDEDVVEINEKFTELLAITNCEERFSLFKNNRSLK 120
Db 61 PSAKGKKNQIGIKILEQPHAVLFVDEDVVEINEKFTELLAITNCEERFSLFKNNRSLK 120
QY 121 GLQIDVGCVPVKQLRSGBEKFPGVVRFRGPLLAERTVSGIFPGVELLEEGRGQGTGVY 180
Db 121 GLQIDVGCVPVKQLRSGBEKFPGVVRFRGPLLAERTVSGIFPGVELLEEGRGQGTGVY 180
QY 181 QGKQLFQCDCEGFFVALDKLELIEDDTALESDDYAGPGDTMQVELPPLLEINRSVSLKGE 240
Db 181 QGKQLFQCDCEGFFVALDKLELIEDDTALESDDYAGPGDTMQVELPPLLEINRSVSLKGE 240
QY 241 TIESGTVIFCDVLPKESIGYPVGVDMNPICGNWDGRDPGVLCSFACVESTILLHINDII 300
Db 241 TIESGTVIFCDVLPKESIGYPVGVDMNPICGNWDGRDPGVLCSFACVESTILLHINDII 300
QY 301 PESVTQERRPPKAPFMSRGVGDGKSSSHNKPATGSTDPGNRRSELFTYTLNGSSVDSQP 360
Db 301 PESVTQERRPPKAPFMSRGVGDGKSSSHNKPATGSTDPGNRRSELFTYTLNGSSVDSQP 360
QY 361 QSKSKNTWYIDEVADPAKSLTEISTDFDRSSPPLOPPPVNSLTNENRPHSLTKMP 420
Db 361 QSKSKNTWYIDEVADPAKSLTEISTDFDRSSPPLOPPPVNSLTNENRPHSLTKMP 420
QY 421 NTNGSIGHSPLSLSAQSVMEELNTAPQESPLAMPNGSHGLEVGSVLAEVKENPPFVGV 480
Db 421 NTNGSIGHSPLSLSAQSVMEELNTAPQESPLAMPNGSHGLEVGSVLAEVKENPPFVGV 480
QY 481 IRWIGQPPGLNEVLAGLEDEBACCTDGTFRGTRYFTCALKKALFVKLKSRRPDSRFAS 540
Db 481 IRWIGQPPGLNEVLAGLEDEBACCTDGTFRGTRYFTCALKKALFVKLKSRRPDSRFAS 540
QY 541 LQPVSNQIERCNSLAFAGGYLSEVVEENTPPKMEKEGLEIMIGKKGIGQHYNVSCYLDSTL 600
Db 541 LQPVSNQIERCNSLAFAGGYLSEVVEENTPPKMEKEGLEIMIGKKGIGQHYNVSCYLDSTL 600
QY 601 FCLFAFSSVLDVTLRPKEKNDVEYVYSETQELLRTTEIVNPLRIYGVVCAATKMKRLKILE 660
Db 601 FCLFAFSSVLDVTLRPKEKNDVEYVYSETQELLRTTEIVNPLRIYGVVCAATKMKRLKILE 660
QY 661 KVEAASGTSBEKOPPEEFNLIFPHILRVEPLLKIRSAGQKQVDCYFYQIFMEKNEKGV 720
Db 661 KVEAASGTSBEKOPPEEFNLIFPHILRVEPLLKIRSAGQKQVDCYFYQIFMEKNEKGV 720
QY 721 PTIQQLLEWSFINSLKFAEAPSCIIIQMPRFGKDFKLFKIFPSPLELNITDLEDTPRQ 780
Db 721 PTIQQLLEWSFINSLKFAEAPSCIIIQMPRFGKDFKLFKIFPSPLELNITDLEDTPRQ 780
QY 781 CRI CGGLAMEYECRECYDDPDISAGIKI KQFCKTNTQVHLHPKRLNHNKYNPVSLPKDLPW 840
Db 781 CRI CGGLAMEYECRECYDDPDISAGIKI KQFCKTNTQVHLHPKRLNHNKYNPVSLPKDLPW 840
QY 841 DWRHGCIPCONNELFAVLCIETSHVAVFKYKGDSDSAMLFPDSDMADRGQGNIPQVT 900
Db 841 DWRHGCIPCONNELFAVLCIETSHVAVFKYKGDSDSAMLFPDSDMADRGQGNIPQVT 900
QY 901 PCPEVGEYLNKMSLEDLHSLDSRRIOGCARRLLCDAYMCMYQSTWMSLYK 949
Db 901 PCPEVGEYLNKMSLEDLHSLDSRRIOGCARRLLCDAYMCMYQSTWMSLYK 949
```

## RESULT 2

```
US-09-488-725A-2399
; Sequence 2399, Application US/09488725A
; GENERAL INFORMATION:
; APPLICANT: Hyseq Inc
; TITLE OF INVENTION: Novel Nucleic Acid and Polypeptides
; FILE REFERENCE: 784PLPCT
; CURRENT APPLICATION NUMBER: US/09/488,725A
; CURRENT FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US/09/488,725
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: US09/598,042
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: US09/620,312
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: US09/653,450
; PRIOR FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: US09/662,191
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: US09/693,036
; PRIOR FILING DATE: 2000-10-19
; PRIOR APPLICATION NUMBER: US09/727,344
; PRIOR FILING DATE: 2000-11-29
; NUMBER OF SEQ ID NOS: 7144
; SOFTWARE: pt_FL_genes_b Versions 1.0
; SEQ ID NO 2399
; LENGTH: 686
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-488-725A-2399
```

Query Match 63.9%; Score 606; DB 18; Length 685;  
Best Local Similarity 100.0%; Pred. No. 0;

```
Matches 606; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 344 RSELYTYTLNGSSVDSQPSQSKNTWYIDEVADPAKSLTEISTDFDRSSPPLOPPPVNSL 403
Db 80 RSELYTYTLNGSSVDSQPSQSKNTWYIDEVADPAKSLTEISTDFDRSSPPLOPPPVNSL 139
QY 404 TTENRPHSLPESLTQKPNNTNGSIGHSPLSLSAQSVMEELNTAPQESPLAMPNGSHGL 463
Db 140 TTENRPHSLPESLTQKPNNTNGSIGHSPLSLSAQSVMEELNTAPQESPLAMPNGSHGL 199
QY 464 EYVGSVLAEVKENPPFVGVIRWIGQPPGLNEVLAGLEDEBACCTDGTFRGTRYFTCALKK 523
Db 200 EYVGSVLAEVKENPPFVGVIRWIGQPPGLNEVLAGLEDEBACCTDGTFRGTRYFTCALKK 259
QY 524 ALFVKLKSRRPDSRFASLQPVSNQIERCNSLAFAGGYLSEVVEENTPPKMEKEGLEIMIGK 583
Db 260 ALFVKLKSRRPDSRFASLQPVSNQIERCNSLAFAGGYLSEVVEENTPPKMEKEGLEIMIGK 319
QY 584 KKGIGQHYNVSCYLDSTLFCFLFAFSSVLDVTLRPKEKNDVEYVYSETQELLRTTEIVNPLRI 643
Db 320 KKGIGQHYNVSCYLDSTLFCFLFAFSSVLDVTLRPKEKNDVEYVYSETQELLRTTEIVNPLRI 379
QY 644 YGVVCAATKMKRLKILEKVEAASGTSBEKOPPEEFNLIFPHILRVEPLLKIRSAGQKQV 703
Db 380 YGVVCAATKMKRLKILEKVEAASGTSBEKOPPEEFNLIFPHILRVEPLLKIRSAGQKQV 439
QY 704 DCYFYQIFMEKNEKGVVPTIQQLLEWSFINSLKFAEAPSCIIIQMPRFGKDFKLFKIF 763
Db 440 DCYFYQIFMEKNEKGVVPTIQQLLEWSFINSLKFAEAPSCIIIQMPRFGKDFKLFKIF 499
QY 764 PSELELNITDLEDTPRQCRICGGLAMEYECRECYDDPDISAGIKI KQFCKTNTQVHLHPKRL 823
Db 500 PSELELNITDLEDTPRQCRICGGLAMEYECRECYDDPDISAGIKI KQFCKTNTQVHLHPKRL 559
QY 824 LNHKNYPVSLPKDLPDSDMADRGQGNIPQVT 883
```

Db 560 LNHNKPNVSLPKDLPDWDRHGCIPQNMELFAVLCTIETSHYVAFVKYKDDSAWLFDFS 619  
Qy 884 MADRDGGQNGFNIPQVTPCPVEGYLKMSLEDLHSLDSRRIOGCARRLLCDAYMCWYQSP 943  
Db 620 MADRDGGQNGFNIPQVTPCPVEGYLKMSLEDLHSLDSRRIOGCARRLLCDAYMCWYQSP 679  
Qy 944 TMSLYK 949  
Db 680 TMSLYK 685  
RESULT 3  
US-10-258-898A-2399  
; Sequence 2399, Application US/10258898A  
; GENERAL INFORMATION:  
; APPLICANT: Hyseq Inc  
; TITLE OF INVENTION: Novel Nucleic Acid and Polypeptides  
; FILE REFERENCE: 784FLPCT  
; CURRENT APPLICATION NUMBER: US/10/258,898A  
; CURRENT FILING DATE: 2002-10-29  
; PRIOR APPLICATION NUMBER: US/09/488,725  
; PRIOR FILING DATE: 2000-01-21  
; PRIOR APPLICATION NUMBER: US09/552,317  
; PRIOR FILING DATE: 2000-04-25  
; PRIOR APPLICATION NUMBER: US09/598,042  
; PRIOR FILING DATE: 2000-06-20  
; PRIOR APPLICATION NUMBER: US09/620,312  
; PRIOR FILING DATE: 2000-07-19  
; PRIOR APPLICATION NUMBER: US09/653,450  
; PRIOR FILING DATE: 2000-08-31  
; PRIOR APPLICATION NUMBER: US09/662,191  
; PRIOR FILING DATE: 2000-09-14  
; PRIOR APPLICATION NUMBER: US09/693,036  
; PRIOR FILING DATE: 2000-10-19  
; PRIOR APPLICATION NUMBER: US09/727,344  
; PRIOR FILING DATE: 2000-11-29  
; NUMBER OF SEQ ID NOS: 7143  
; SOFTWARE: pt\_FL\_genes\_b Versions 1.0  
; SEQ ID NO 2399  
; LENGTH: 685  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-258-898A-2399

Query Match 63.9%; Score 606; DB 28; Length 685;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 606; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 344 RSELFYTLNGSSVDSQPSQSKNTWYIDEVAEDPAKSLTEISTDFDRSSPPLQPPVNSL 403  
Db 80 RSELFYTLNGSSVDSQPSQSKNTWYIDEVAEDPAKSLTEISTDFDRSSPPLQPPVNSL 139  
Qy 404 TTENRFHSLPFLSLTKMPNTNGSIGHSPLSLSAQSVMEELNTPAQVQSPPLAMPNGSHGL 463  
Db 140 TTENRFHSLPFLSLTKMPNTNGSIGHSPLSLSAQSVMEELNTPAQVQSPPLAMPNGSHGL 199  
Qy 464 EVGSLAEVKENPPFYGVIRWIGQPGNLVLAAGLEDEACAGTDTGTFGRTRYFTCALKK 523  
Db 200 EVGSLAEVKENPPFYGVIRWIGQPGNLVLAAGLEDEACAGTDTGTFGRTRYFTCALKK 259  
Qy 524 ALFVKLKSCRDPDSRSLQPVSNQTERCNSLAFGGYLSVEVENTPPKMEKGLTMIWK 583  
Db 260 ALFVKLKSCRDPDSRSLQPVSNQTERCNSLAFGGYLSVEVENTPPKMEKGLTMIWK 319  
Qy 584 KKGIGCHYNSCYLDSTLFCFLAFSSVLDTVLLRPKEKNVDEYVYSETQELLRTIENVPLRI 643  
Db 320 KKGIGCHYNSCYLDSTLFCFLAFSSVLDTVLLRPKEKNVDEYVYSETQELLRTIENVPLRI 379  
Qy 644 YGYVCATKIMLKURKILEKVEAASGFTSEBKPEEFLNLFHILRVEPLLLKIRSAQKQV 703  
Db 380 YGYVCATKIMLKURKILEKVEAASGFTSEBKPEEFLNLFHILRVEPLLLKIRSAQKQV 439  
Qy 704 DCYFYQIFMEKNEKVGVPITIQQLLEWSFINSNLKFAEAFSCLLIQMPRFKDFKLFKIF 763

Db 440 DCYFYQIFMEKNEKVGVPITIQQLLEWSFINSNLKFAEAFSCLLIQMPRFKDFKLFKIF 499  
Qy 764 PSLELNITDLEETPRQCRICGGLAMYECRECYDDPD1SAGKIKQFCKTCNTQVHLHPKR 823  
Db 500 PSLELNITDLEETPRQCRICGGLAMYECRECYDDPD1SAGKIKQFCKTCNTQVHLHPKR 559  
Qy 824 LNHNKPNVSLPKDLPDWDRHGCIPQNMELFAVLCTIETSHYVAFVKYKDDSAWLFDFS 883  
Db 560 LNHNKPNVSLPKDLPDWDRHGCIPQNMELFAVLCTIETSHYVAFVKYKDDSAWLFDFS 619  
Qy 884 MADRDGGQNGFNIPQVTPCPVEGYLKMSLEDLHSLDSRRIOGCARRLLCDAYMCWYQSP 943  
Db 620 MADRDGGQNGFNIPQVTPCPVEGYLKMSLEDLHSLDSRRIOGCARRLLCDAYMCWYQSP 679  
Qy 944 TMSLYK 949  
Db 680 TMSLYK 685  
RESULT 4  
US-10-286-897-2399  
; Sequence 2399, Application US/10286897  
; GENERAL INFORMATION:  
; APPLICANT: Hyseq Inc  
; TITLE OF INVENTION: Novel Nucleic Acid and Polypeptides  
; FILE REFERENCE: 784FLPCT  
; CURRENT APPLICATION NUMBER: US/10/286,897  
; CURRENT FILING DATE: 2002-11-01  
; PRIOR APPLICATION NUMBER: US/09/488,725  
; PRIOR FILING DATE: 2000-01-21  
; PRIOR APPLICATION NUMBER: US/09/552,317  
; PRIOR FILING DATE: 2000-04-25  
; PRIOR APPLICATION NUMBER: US/09/598,042  
; PRIOR FILING DATE: 2000-06-20  
; PRIOR APPLICATION NUMBER: US/09/620,312  
; PRIOR FILING DATE: 2000-07-19  
; PRIOR APPLICATION NUMBER: US/09/653,450  
; PRIOR FILING DATE: 2000-08-31  
; PRIOR APPLICATION NUMBER: US/09/662,191  
; PRIOR FILING DATE: 2000-09-14  
; PRIOR APPLICATION NUMBER: US/09/693,036  
; PRIOR FILING DATE: 2000-10-19  
; PRIOR APPLICATION NUMBER: US/09/727,344  
; PRIOR FILING DATE: 2000-11-29  
; NUMBER OF SEQ ID NOS: 7143  
; SOFTWARE: pt\_FL\_genes\_b Versions 1.0  
; SEQ ID NO 2399  
; LENGTH: 685  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-286-897-2399

Query Match 63.9%; Score 606; DB 28; Length 685;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 606; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 344 RSELFYTLNGSSVDSQPSQSKNTWYIDEVAEDPAKSLTEISTDFDRSSPPLQPPVNSL 403  
Db 80 RSELFYTLNGSSVDSQPSQSKNTWYIDEVAEDPAKSLTEISTDFDRSSPPLQPPVNSL 139  
Qy 404 TTENRFHSLPFLSLTKMPNTNGSIGHSPLSLSAQSVMEELNTPAQVQSPPLAMPNGSHGL 463  
Db 140 TTENRFHSLPFLSLTKMPNTNGSIGHSPLSLSAQSVMEELNTPAQVQSPPLAMPNGSHGL 199  
Qy 464 EVGSLAEVKENPPFYGVIRWIGQPGNLVLAAGLEDEACAGTDTGTFGRTRYFTCALKK 523  
Db 200 EVGSLAEVKENPPFYGVIRWIGQPGNLVLAAGLEDEACAGTDTGTFGRTRYFTCALKK 259  
Qy 524 ALFVKLKSCRDPDSRSLQPVSNQTERCNSLAFGGYLSVEVENTPPKMEKGLTMIWK 583  
Db 260 ALFVKLKSCRDPDSRSLQPVSNQTERCNSLAFGGYLSVEVENTPPKMEKGLTMIWK 319

QY 584 KKGIOGHVNSCYLDSTLFCFAFSSVLDTVLLRPKEKNDVEYYSQTELLRTEIYNPLRI 643  
DB 320 KKGIOGHVNSCYLDSTLFCFAFSSVLDTVLLRPKEKNDVEYYSQTELLRTEIYNPLRI 379  
QY 644 YGYVCATKIMLKRLKILEKVEAASGFTSEKOPPEEFNLILFHILRVEPLLKIRSGQKQV 703  
DB 380 YGYVCATKIMLKRLKILEKVEAASGFTSEKOPPEEFNLILFHILRVEPLLKIRSGQKQV 439  
QY 704 DCYFYQIFMEKNEKVGVPITIOQLLEWSPINSNLKFAEAPSCIIQMPRFGKDFKLFKKIF 763  
DB 440 DCYFYQIFMEKNEKVGVPITIOQLLEWSPINSNLKFAEAPSCIIQMPRFGKDFKLFKKIF 499  
QY 764 PSLEINITDLEDTPRQCRICGGLAMYECCYDDPDISAGKIKOFCCTCNTQVHLHPR 823  
DB 500 PSLEINITDLEDTPRQCRICGGLAMYECCYDDPDISAGKIKOFCCTCNTQVHLHPR 559  
QY 824 LNHKYNPVSPLKDLDPDWRHGCIPQNMELFAVLCIETSHYVAFVKYKGDSDSAMLFFDS 883  
DB 560 LNHKYNPVSPLKDLDPDWRHGCIPQNMELFAVLCIETSHYVAFVKYKGDSDSAMLFFDS 619  
QY 884 MADRDGGGNGFNIPQVTPCPEVGEYVKMSLEDLHSLDSRRIOGCARRLLCDAYMCWQSP 943  
DB 620 MADRDGGGNGFNIPQVTPCPEVGEYVKMSLEDLHSLDSRRIOGCARRLLCDAYMCWQSP 679  
QY 944 TMSLYK 949  
DB 680 TMSLYK 685

## RESULT 5

US-09-786-797B-9  
; Sequence 9, Application US/09786797B  
; GENERAL INFORMATION:  
; APPLICANT: INCYTE PHARMACEUTICALS, INC.  
; APPLICANT: LAL, Preeti  
; APPLICANT: TANG, Y. Tom  
; APPLICANT: YUE, Henry  
; APPLICANT: HILLMAN, Jennifer L.  
; APPLICANT: BANDMAN, Olga  
; APPLICANT: CORLEY, Neil C.  
; APPLICANT: GUEGLER, Karl J.  
; APPLICANT: PATTERSON, Chandra  
; APPLICANT: AZIMZAI, Yalda  
; APPLICANT: BAUGHN, Mariah R.  
; TITLE OF INVENTION: HUMAN CYTOSKELETON ASSOCIATED PROTEINS  
; FILE REFERENCE: PF-0594 PCT  
; CURRENT APPLICATION NUMBER: US/09/786,797B  
; CURRENT FILING DATE: 2002-08-26  
; PRIOR APPLICATION NUMBER: 09/156,470; unassigned; 60/131,321  
; PRIOR FILING DATE: 1998-09-18; 1998-09-18; 1999-04-27  
; NUMBER OF SEQ ID NOS: 32  
; SOFTWARE: PERL Program  
; SEQ ID NO 9  
; LENGTH: 731  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc.feature  
; OTHER INFORMATION: Incyte ID No: 2363327  
US-09-786-797B-9

Query Match 63.9%; Score 606; DB 22; Length 731;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 606; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 344 RSELFYTLNGSVDSQPSKNTWYIDEVADPAKSTLSTDFDRSSPPLQPPPVNSL 403  
DB 126 RSELFYTLNGSVDSQPSKNTWYIDEVADPAKSTLSTDFDRSSPPLQPPPVNSL 185  
QY 404 TTRNRFHSLPFLTKMPNTNGSIGHSPLSLSAQSVMEELNTPAQESPPPLAMPNGSHGL 463  
DB 186 TTRNRFHSLPFLTKMPNTNGSIGHSPLSLSAQSVMEELNTPAQESPPPLAMPNGSHGL 245

QY 464 EVGSLAEVKENPPFYGVIRWIGQPPGLNEVLAGELEDEACAGCTDGTFRGTRYFTCALKK 523  
DB 246 EVGSLAEVKENPPFYGVIRWIGQPPGLNEVLAGELEDEACAGCTDGTFRGTRYFTCALKK 305  
QY 524 ALFVKLKSRRDPSRSLQPVSNQIERCNSLAFGGYLSVUVEENTPPKMEKEGLEIMIGK 583  
DB 306 ALFVKLKSRRDPSRSLQPVSNQIERCNSLAFGGYLSVUVEENTPPKMEKEGLEIMIGK 365  
QY 584 KKGIOGHVNSCYLDSTLFCFAFSSVLDTVLLRPKEKNDVEYYSQTELLRTEIYNPLRI 643  
DB 366 KKGIOGHVNSCYLDSTLFCFAFSSVLDTVLLRPKEKNDVEYYSQTELLRTEIYNPLRI 425  
QY 644 YGYVCATKIMLKRLKILEKVEAASGFTSEKOPPEEFNLILFHILRVEPLLKIRSGQKQV 703  
DB 426 YGYVCATKIMLKRLKILEKVEAASGFTSEKOPPEEFNLILFHILRVEPLLKIRSGQKQV 485  
QY 704 DCYFYQIFMEKNEKVGVPITIOQLLEWSPINSNLKFAEAPSCIIQMPRFGKDFKLFKKIF 763  
DB 486 DCYFYQIFMEKNEKVGVPITIOQLLEWSPINSNLKFAEAPSCIIQMPRFGKDFKLFKKIF 545  
QY 764 PSLEINITDLEDTPRQCRICGGLAMYECCYDDPDISAGKIKOFCCTCNTQVHLHPR 823  
DB 546 PSLEINITDLEDTPRQCRICGGLAMYECCYDDPDISAGKIKOFCCTCNTQVHLHPR 605  
QY 824 LNHKYNPVSPLKDLDPDWRHGCIPQNMELFAVLCIETSHYVAFVKYKGDSDSAMLFFDS 883  
DB 606 LNHKYNPVSPLKDLDPDWRHGCIPQNMELFAVLCIETSHYVAFVKYKGDSDSAMLFFDS 665  
QY 884 MADRDGGGNGFNIPQVTPCPEVGEYVKMSLEDLHSLDSRRIOGCARRLLCDAYMCWQSP 943  
DB 666 MADRDGGGNGFNIPQVTPCPEVGEYVKMSLEDLHSLDSRRIOGCARRLLCDAYMCWQSP 725  
QY 944 TMSLYK 949  
DB 726 TMSLYK 731

## RESULT 6

US-10-921-707-9  
; Sequence 9, Application US/10921707  
; GENERAL INFORMATION:  
; APPLICANT: INCYTE PHARMACEUTICALS, INC.  
; APPLICANT: LAL, Preeti  
; APPLICANT: TANG, Y. Tom  
; APPLICANT: YUE, Henry  
; APPLICANT: HILLMAN, Jennifer L.  
; APPLICANT: BANDMAN, Olga  
; APPLICANT: CORLEY, Neil C.  
; APPLICANT: GUEGLER, Karl J.  
; APPLICANT: PATTERSON, Chandra  
; APPLICANT: AZIMZAI, Yalda  
; APPLICANT: BAUGHN, Mariah R.  
; TITLE OF INVENTION: HUMAN CYTOSKELETON ASSOCIATED PROTEINS  
; FILE REFERENCE: PF-0594 PCT  
; CURRENT APPLICATION NUMBER: US/10/921,707  
; CURRENT FILING DATE: 2004-08-19  
; PRIOR APPLICATION NUMBER: US/09/786,797  
; PRIOR FILING DATE: 2001-06-25  
; PRIOR APPLICATION NUMBER: 09/156,470; unassigned; 60/131,321  
; PRIOR FILING DATE: 1998-09-18; 1998-09-18; 1999-04-27  
; NUMBER OF SEQ ID NOS: 32  
; SOFTWARE: PERL Program  
; SEQ ID NO 9  
; LENGTH: 731  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc.feature  
; OTHER INFORMATION: Incyte ID No: 2363327  
US-10-921-707-9

Query Match 63.9%; Score 606; DB 35; Length 731;  
Best Local Similarity 100.0%; Pred. No. 0;



404 TTNRFHSLPFLSLTKMPNTNGSIHSPLSLSAQSVMEELNAPVQESPLAMPNGSHGL 463  
194 TTNRFHSLPFLSLTKMPNTNGSIHSPLSLSAQSVMEELNAPVQESPLAMPNGSHGL 253  
464 EVGSLAEVKNPPFYGVIRWIGQPPGLNEVLAGELEDECACTDGTFRGTRYFTCALKK 523  
254 EVGSLAEVKNPPFYGVIRWIGQPPGLNEVLAGELEDECACTDGTFRGTRYFTCALKK 313  
524 ALFVKLSKCRDPSRPFASIQPVSNQIRCNLSAFGGYLSVVEENTPPKMEKEGLEIMIGK 583  
314 ALFVKLSKCRDPSRPFASIQPVSNQIRCNLSAFGGYLSVVEENTPPKMEKEGLEIMIGK 373  
584 KKGIOGHVNSCYLDSTLFCFLFAFSSVLDTVLLRPKEKNDVEYYSQTELLRTEIVNPLRI 643  
374 KKGIOGHVNSCYLDSTLFCFLFAFSSVLDTVLLRPKEKNDVEYYSQTELLRTEIVNPLRI 433  
644 YGYVCATKIMKRLKILEKVEAASGTSSEKDPPEEFLNLFHHLRVEPLLKIRSAGQKVQ 703  
434 YGYVCATKIMKRLKILEKVEAASGTSSEKDPPEEFLNLFHHLRVEPLLKIRSAGQKVQ 493  
704 DCYFQIEMEKNEKVGVPVPTIOQLLEWSFINSNLKFAEAPSCLIQMPRFGKDFKPKKIF 763  
494 DCYFQIEMEKNEKVGVPVPTIOQLLEWSFINSNLKFAEAPSCLIQMPRFGKDFKPKKIF 553  
764 PSLEINITDLEDTPRQCRI CGGLAMYECRECYDDPDISAGKIKQFCCKTQVHLHPKR 823  
554 PSLEINITDLEDTPRQCRI CGGLAMYECRECYDDPDISAGKIKQFCCKTQVHLHPKR 613  
824 LNHNKYNPVSPLPKDLPDWRHGCIPQNNMELFAVLCIETSHYVAFVKYKODSAMLFFDS 883  
614 LNHNKYNPVSPLPKDLPDWRHGCIPQNNMELFAVLCIETSHYVAFVKYKODSAMLFFDS 673  
884 MADRDGGQGNFIPQVTPCPEVGEYVKMSLEDLHLSRRIOGCARRLLCDAYMCYQSP 943  
674 MADRDGGQGNFIPQVTPCPEVGEYVKMSLEDLHLSRRIOGCARRLLCDAYMCYQSP 733  
944 TMSLYK 949  
734 TMSLYK 739

## RESULT 9

US-09-764-902-1743  
; Sequence 1743, Application US/09764902  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PT213  
; CURRENT APPLICATION NUMBER: US/09/764,902  
; CURRENT FILING DATE: 2001-01-17  
; Prior application data removed - refer to PALM or file wrapper  
; NUMBER OF SEQ ID NOS: 2318  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1743  
; LENGTH: 739  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-764-902-1743

Query Match 63.9%; Score 606; DB 22; Length 739;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 606; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

344 RSELYTLNGSSVDSQPSKNTWYIDEVAEDPAKSLTEISTDFDRSSPPLQPPPVNSL 403  
134 RSELYTLNGSSVDSQPSKNTWYIDEVAEDPAKSLTEISTDFDRSSPPLQPPPVNSL 193  
404 TTNRFHSLPFLSLTKMPNTNGSIHSPLSLSAQSVMEELNAPVQESPLAMPNGSHGL 463  
194 TTNRFHSLPFLSLTKMPNTNGSIHSPLSLSAQSVMEELNAPVQESPLAMPNGSHGL 253  
464 EVGSLAEVKNPPFYGVIRWIGQPPGLNEVLAGELEDECACTDGTFRGTRYFTCALKK 523

254 EVGSLAEVKNPPFYGVIRWIGQPPGLNEVLAGELEDECACTDGTFRGTRYFTCALKK 313  
524 ALFVKLSKCRDPSRPFASIQPVSNQIRCNLSAFGGYLSVVEENTPPKMEKEGLEIMIGK 583  
314 ALFVKLSKCRDPSRPFASIQPVSNQIRCNLSAFGGYLSVVEENTPPKMEKEGLEIMIGK 373  
584 KKGIOGHVNSCYLDSTLFCFLFAFSSVLDTVLLRPKEKNDVEYYSQTELLRTEIVNPLRI 643  
374 KKGIOGHVNSCYLDSTLFCFLFAFSSVLDTVLLRPKEKNDVEYYSQTELLRTEIVNPLRI 433  
644 YGYVCATKIMKRLKILEKVEAASGTSSEKDPPEEFLNLFHHLRVEPLLKIRSAGQKVQ 703  
434 YGYVCATKIMKRLKILEKVEAASGTSSEKDPPEEFLNLFHHLRVEPLLKIRSAGQKVQ 493  
704 DCYFQIEMEKNEKVGVPVPTIOQLLEWSFINSNLKFAEAPSCLIQMPRFGKDFKPKKIF 763  
494 DCYFQIEMEKNEKVGVPVPTIOQLLEWSFINSNLKFAEAPSCLIQMPRFGKDFKPKKIF 553  
764 PSLEINITDLEDTPRQCRI CGGLAMYECRECYDDPDISAGKIKQFCCKTQVHLHPKR 823  
554 PSLEINITDLEDTPRQCRI CGGLAMYECRECYDDPDISAGKIKQFCCKTQVHLHPKR 613  
824 LNHNKYNPVSPLPKDLPDWRHGCIPQNNMELFAVLCIETSHYVAFVKYKODSAMLFFDS 883  
614 LNHNKYNPVSPLPKDLPDWRHGCIPQNNMELFAVLCIETSHYVAFVKYKODSAMLFFDS 673  
884 MADRDGGQGNFIPQVTPCPEVGEYVKMSLEDLHLSRRIOGCARRLLCDAYMCYQSP 943  
674 MADRDGGQGNFIPQVTPCPEVGEYVKMSLEDLHLSRRIOGCARRLLCDAYMCYQSP 733  
944 TMSLYK 949  
734 TMSLYK 739

## RESULT 10

PCT-US02-27777-137  
; Sequence 137, Application PC/TUS0227777  
; GENERAL INFORMATION:  
; APPLICANT: diadexus, Inc.  
; APPLICANT: Sun, Yongming  
; APPLICANT: Liu, Chenghua  
; APPLICANT: Salceda, Susana  
; TITLE OF INVENTION: Compositions and Methods Relating to Breast Specific Genes and Pr  
; FILE REFERENCE: DEX-0346  
; CURRENT APPLICATION NUMBER: PCT/US02/27777  
; CURRENT FILING DATE: 2002-08-29  
; PRIOR APPLICATION NUMBER: US 60/316,306  
; PRIOR FILING DATE: 2001-08-31  
; NUMBER OF SEQ ID NOS: 170  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 137  
; LENGTH: 953  
; TYPE: PRT  
; ORGANISM: Homo sapien  
PCT-US02-27777-137

Query Match 63.9%; Score 606; DB 1; Length 953;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 606; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

344 RSELYTLNGSSVDSQPSKNTWYIDEVAEDPAKSLTEISTDFDRSSPPLQPPPVNSL 403  
348 RSELYTLNGSSVDSQPSKNTWYIDEVAEDPAKSLTEISTDFDRSSPPLQPPPVNSL 407  
404 TTNRFHSLPFLSLTKMPNTNGSIHSPLSLSAQSVMEELNAPVQESPLAMPNGSHGL 463  
408 TTNRFHSLPFLSLTKMPNTNGSIHSPLSLSAQSVMEELNAPVQESPLAMPNGSHGL 467  
464 EVGSLAEVKNPPFYGVIRWIGQPPGLNEVLAGELEDECACTDGTFRGTRYFTCALKK 523  
468 EVGSLAEVKNPPFYGVIRWIGQPPGLNEVLAGELEDECACTDGTFRGTRYFTCALKK 527

QY 524 ALFVKLSKCRPSRPFASLQPVSNQIERCNSLAFGGYLSVVEENTPPKMEKEGLEIMIGK 583  
DB 528 ALFVKLSKCRPSRPFASLQPVSNQIERCNSLAFGGYLSVVEENTPPKMEKEGLEIMIGK 587  
QY 584 KKGIOGHYNSCYLDSTLFCFLFAPSSVLDTVLLRPKEKNDVEYYSETQELLRTIENVPLRI 643  
DB 588 KKGIOGHYNSCYLDSTLFCFLFAPSSVLDTVLLRPKEKNDVEYYSETQELLRTIENVPLRI 647  
QY 644 YGYVCATKIMKRLKILEKVEAASGFTSEBKDPPEEFNLILFHILRLRVEPLLKIRSAGOKVQ 703  
DB 648 YGYVCATKIMKRLKILEKVEAASGFTSEBKDPPEEFNLILFHILRLRVEPLLKIRSAGOKVQ 707  
QY 704 DCYFYQIEMKNEKVGVPITIQQLLEWSFINSNLKFAEAPSCLIIOMPREGKDFKLPFKKIF 763  
DB 708 DCYFYQIEMKNEKVGVPITIQQLLEWSFINSNLKFAEAPSCLIIOMPREGKDFKLPFKKIF 767  
QY 764 PSLELNITDLLEDTTPQRCRI CGGLAMYECRECYDDPDISAGIKQFCCTCNTQVHLHPKR 823  
DB 768 PSLELNITDLLEDTTPQRCRI CGGLAMYECRECYDDPDISAGIKQFCCTCNTQVHLHPKR 827  
QY 824 LNHKNPVSPLPKDLPDWDWRHGCIPCONMELFAVLCIETSHYVAFVKYKDDSAWLFFDS 883  
DB 828 LNHKNPVSPLPKDLPDWDWRHGCIPCONMELFAVLCIETSHYVAFVKYKDDSAWLFFDS 887  
QY 884 MADRDGGQNGFNIPQVTPCPEVGEYKMSLEDLHSLDSRRIOGCARRLLCDAYMCYQSP 943  
DB 888 MADRDGGQNGFNIPQVTPCPEVGEYKMSLEDLHSLDSRRIOGCARRLLCDAYMCYQSP 947  
QY 944 TMSLYK 949  
DB 948 TMSLYK 953

RESULT 11  
PCT-US02-27777A-137  
; Sequence 137, Application PC/TUS0227777A  
; GENERAL INFORMATION:  
; APPLICANT: diadexus, Inc.  
; APPLICANT: Sun, Yongming  
; APPLICANT: Liu, Chenghua  
; APPLICANT: Salceda, Susana  
; TITLE OF INVENTION: Compositions and Methods related to Breast Specific Genes and Pro  
; FILE REFERENCE: DEX-0346  
; CURRENT APPLICATION NUMBER: PCT/US02/27777A  
; CURRENT FILING DATE: 2002-10-24  
; PRIOR APPLICATION NUMBER: US 60/316,307  
; PRIOR FILING DATE: 2001-08-31  
; NUMBER OF SEQ ID NOS: 170  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 137  
; LENGTH: 953  
; TYPE: PRT  
; ORGANISM: Homo sapien  
PCT-US02-27777A-137

Query Match 63.9%; Score 606; DB 1; Length 953;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 606; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 344 RSELYTLNGSSVDSQPSKSKNTWYIDEVAEDPAKSLTEISTDFDRSSPPLQPPPVNSL 403  
DB 348 RSELYTLNGSSVDSQPSKSKNTWYIDEVAEDPAKSLTEISTDFDRSSPPLQPPPVNSL 407  
QY 404 TTENRFHSLPFLSLTKMPNTNGSIGHSPLSLSAQSVMEELNAPVQESPPPLAMPNGSHGL 463  
DB 408 TTENRFHSLPFLSLTKMPNTNGSIGHSPLSLSAQSVMEELNAPVQESPPPLAMPNGSHGL 467  
QY 464 EVGSLAEVKENPPFFYGVIRWIGQPPGLNEVLAGELEDECACTDGTGTRGTYFTCALKK 523  
DB 468 EVGSLAEVKENPPFFYGVIRWIGQPPGLNEVLAGELEDECACTDGTGTRGTYFTCALKK 527  
QY 524 ALFVKLSKCRPSRPFASLQPVSNQIERCNSLAFGGYLSVVEENTPPKMEKEGLEIMIGK 583

DB 528 ALFVKLSKCRPSRPFASLQPVSNQIERCNSLAFGGYLSVVEENTPPKMEKEGLEIMIGK 587  
QY 584 KKGIOGHYNSCYLDSTLFCFLFAPSSVLDTVLLRPKEKNDVEYYSETQELLRTIENVPLRI 643  
DB 588 KKGIOGHYNSCYLDSTLFCFLFAPSSVLDTVLLRPKEKNDVEYYSETQELLRTIENVPLRI 647  
QY 644 YGYVCATKIMKRLKILEKVEAASGFTSEBKDPPEEFNLILFHILRLRVEPLLKIRSAGOKVQ 703  
DB 648 YGYVCATKIMKRLKILEKVEAASGFTSEBKDPPEEFNLILFHILRLRVEPLLKIRSAGOKVQ 707  
QY 704 DCYFYQIEMKNEKVGVPITIQQLLEWSFINSNLKFAEAPSCLIIOMPREGKDFKLPFKKIF 763  
DB 708 DCYFYQIEMKNEKVGVPITIQQLLEWSFINSNLKFAEAPSCLIIOMPREGKDFKLPFKKIF 767  
QY 764 PSLELNITDLLEDTTPQRCRI CGGLAMYECRECYDDPDISAGIKQFCCTCNTQVHLHPKR 823  
DB 768 PSLELNITDLLEDTTPQRCRI CGGLAMYECRECYDDPDISAGIKQFCCTCNTQVHLHPKR 827  
QY 824 LNHKNPVSPLPKDLPDWDWRHGCIPCONMELFAVLCIETSHYVAFVKYKDDSAWLFFDS 883  
DB 828 LNHKNPVSPLPKDLPDWDWRHGCIPCONMELFAVLCIETSHYVAFVKYKDDSAWLFFDS 887  
QY 884 MADRDGGQNGFNIPQVTPCPEVGEYKMSLEDLHSLDSRRIOGCARRLLCDAYMCYQSP 943  
DB 888 MADRDGGQNGFNIPQVTPCPEVGEYKMSLEDLHSLDSRRIOGCARRLLCDAYMCYQSP 947  
QY 944 TMSLYK 949  
DB 948 TMSLYK 953

RESULT 12  
US-10-170-205E-37570  
; Sequence 37570, Application US/10170205E  
; GENERAL INFORMATION:  
; APPLICANT: ADAMS, Mark  
; TITLE OF INVENTION: DEVICES, SUCH AS ARRAYS, COMPRISED OF HUMAN PROTEINS OR PROTEIN  
; TITLE OF INVENTION: CAPTURE AGENTS, AND USES THEREOF  
; FILE REFERENCE: CL001381  
; CURRENT APPLICATION NUMBER: US/10/170,205E  
; CURRENT FILING DATE: 2002-06-13  
; NUMBER OF SEQ ID NOS: 40312  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 37570  
; LENGTH: 953  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-170-205E-37570

Query Match 63.9%; Score 606; DB 27; Length 953;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 606; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 344 RSELYTLNGSSVDSQPSKSKNTWYIDEVAEDPAKSLTEISTDFDRSSPPLQPPPVNSL 403  
DB 348 RSELYTLNGSSVDSQPSKSKNTWYIDEVAEDPAKSLTEISTDFDRSSPPLQPPPVNSL 407  
QY 404 TTENRFHSLPFLSLTKMPNTNGSIGHSPLSLSAQSVMEELNAPVQESPPPLAMPNGSHGL 463  
DB 408 TTENRFHSLPFLSLTKMPNTNGSIGHSPLSLSAQSVMEELNAPVQESPPPLAMPNGSHGL 467  
QY 464 EVGSLAEVKENPPFFYGVIRWIGQPPGLNEVLAGELEDECACTDGTGTRGTYFTCALKK 523  
DB 468 EVGSLAEVKENPPFFYGVIRWIGQPPGLNEVLAGELEDECACTDGTGTRGTYFTCALKK 527  
QY 524 ALFVKLSKCRPSRPFASLQPVSNQIERCNSLAFGGYLSVVEENTPPKMEKEGLEIMIGK 583  
DB 528 ALFVKLSKCRPSRPFASLQPVSNQIERCNSLAFGGYLSVVEENTPPKMEKEGLEIMIGK 587  
QY 584 KKGIOGHYNSCYLDSTLFCFLFAPSSVLDTVLLRPKEKNDVEYYSETQELLRTIENVPLRI 643  
DB 588 KKGIOGHYNSCYLDSTLFCFLFAPSSVLDTVLLRPKEKNDVEYYSETQELLRTIENVPLRI 647



QY 644 YGYVCATKIMKRLKILEKVEAASGFTSEKDPPEEFLNLFPHILRVEPLLKIRSAGQKVQ 703  
Db 648 YGYVCATKIMKRLKILEKVEAASGFTSEKDPPEEFLNLFPHILRVEPLLKIRSAGQKVQ 707  
QY 704 DCYFQIIFMEKNEKVGVPITIOQLLEWSFINSNLKFAEAPSCLIIOMPRFGKDFLKKIF 763  
Db 708 DCYFQIIFMEKNEKVGVPITIOQLLEWSFINSNLKFAEAPSCLIIOMPRFGKDFLKKIF 767  
QY 764 PSLELNITDLEDTPROCRICGGLAMVRECYDDPDISAGKIKQFCCKTQNTQVHLHPRK 823  
Db 768 PSLELNITDLEDTPROCRICGGLAMVRECYDDPDISAGKIKQFCCKTQNTQVHLHPRK 827  
QY 824 LNHNKPNVSLPKDLPDMDWRHGCIPQNMELFAVLCIETSHYVAFVKYKDDSAWLFDFS 883  
Db 828 LNHNKPNVSLPKDLPDMDWRHGCIPQNMELFAVLCIETSHYVAFVKYKDDSAWLFDFS 887  
QY 884 MADRDGGONGFNIPQVTPCPEVGEYKMSLEDLHSLDSRRIOGCARRLLCDAYMCWYQSP 943  
Db 888 MADRDGGONGFNIPQVTPCPEVGEYKMSLEDLHSLDSRRIOGCARRLLCDAYMCWYQSP 947  
QY 944 TMSLYK 949  
Db 948 TMSLYK 953

RESULT 13  
PCT-US02-14570-4  
; Sequence 4, Application PC/TUS0214570  
; GENERAL INFORMATION:  
; APPLICANT: IMMUNEX CORPORATION  
; APPLICANT: Derry, Jonathan M. J.  
; APPLICANT: Fanslow III, William  
; APPLICANT: Dougall, William C.  
; TITLE OF INVENTION: SCREENING ASSAYS FOR AGONISTS OR ANTAGONISTS OF CD40 SIGNALING  
; FILE REFERENCE: 3198-WO  
; CURRENT APPLICATION NUMBER: PCT/US02/14570  
; CURRENT FILING DATE: 2002-05-07  
; PRIOR APPLICATION NUMBER: 09/851,673  
; PRIOR FILING DATE: 2001-05-08  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 4  
; LENGTH: 956  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
PCT-US02-14570-4

Query Match 63.9%; Score 606; DB 1; Length 956;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 606; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 344 RSELYTLNGSSVDSQPOSQSKNTWYIDEVAEDPAKSLTEISTDFDRSSPPLQPPPVNSL 403  
Db 351 RSELYTLNGSSVDSQPOSQSKNTWYIDEVAEDPAKSLTEISTDFDRSSPPLQPPPVNSL 410  
QY 404 TTNRFHSLPSLTKMPNTNGSIGHSPLSLSAQSVMBELNAPVQESPLAMPNGSHGL 463  
Db 411 TTNRFHSLPSLTKMPNTNGSIGHSPLSLSAQSVMBELNAPVQESPLAMPNGSHGL 470  
QY 464 EVGSLAEVKNPPFYGVIRWIGQPPGLNEVLAGELEDEACAGCTDGTFRGTRYFTCALKK 523  
Db 471 EVGSLAEVKNPPFYGVIRWIGQPPGLNEVLAGELEDEACAGCTDGTFRGTRYFTCALKK 530  
QY 524 ALFVKLSCRPDSRFASLPQVSNQIERCNSLAFGGYLSVEVEENTPPKMEKEGLEIMIGK 583  
Db 531 ALFVKLSCRPDSRFASLPQVSNQIERCNSLAFGGYLSVEVEENTPPKMEKEGLEIMIGK 590  
QY 584 KKGIOGHVNSCYLDSTLFCFLAFSSVLDVLLRPKEKNDVEYSETQELLRTIENVPLRI 643  
Db 591 KKGIOGHVNSCYLDSTLFCFLAFSSVLDVLLRPKEKNDVEYSETQELLRTIENVPLRI 650  
QY 644 YGYVCATKIMKRLKILEKVEAASGFTSEKDPPEEFLNLFPHILRVEPLLKIRSAGQKVQ 703

Db 651 YGYVCATKIMKRLKILEKVEAASGFTSEKDPPEEFLNLFPHILRVEPLLKIRSAGQKVQ 710  
QY 704 DCYFQIIFMEKNEKVGVPITIOQLLEWSFINSNLKFAEAPSCLIIOMPRFGKDFLKKIF 763  
Db 711 DCYFQIIFMEKNEKVGVPITIOQLLEWSFINSNLKFAEAPSCLIIOMPRFGKDFLKKIF 770  
QY 764 PSLELNITDLEDTPROCRICGGLAMVRECYDDPDISAGKIKQFCCKTQNTQVHLHPRK 823  
Db 771 PSLELNITDLEDTPROCRICGGLAMVRECYDDPDISAGKIKQFCCKTQNTQVHLHPRK 830  
QY 824 LNHNKPNVSLPKDLPDMDWRHGCIPQNMELFAVLCIETSHYVAFVKYKDDSAWLFDFS 883  
Db 831 LNHNKPNVSLPKDLPDMDWRHGCIPQNMELFAVLCIETSHYVAFVKYKDDSAWLFDFS 890  
QY 884 MADRDGGONGFNIPQVTPCPEVGEYKMSLEDLHSLDSRRIOGCARRLLCDAYMCWYQSP 943  
Db 891 MADRDGGONGFNIPQVTPCPEVGEYKMSLEDLHSLDSRRIOGCARRLLCDAYMCWYQSP 950  
QY 944 TMSLYK 949  
Db 951 TMSLYK 956

RESULT 14  
US-09-851-673-4  
; Sequence 4, Application US/09851673  
; GENERAL INFORMATION:  
; APPLICANT: Derry, Jonathan  
; APPLICANT: Fanslow, William  
; APPLICANT: Dougall, William  
; TITLE OF INVENTION: SCREENING ASSAYS FOR AGONISTS OR ANTAGONISTS OF CD40 SIGNALING  
; FILE REFERENCE: 3198  
; CURRENT APPLICATION NUMBER: US/09/851,673  
; CURRENT FILING DATE: 2001-05-08  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: Patent in version 3.0  
; SEQ ID NO 4  
; LENGTH: 956  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-851-673-4

Query Match 63.9%; Score 606; DB 23; Length 956;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 606; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 344 RSELYTLNGSSVDSQPOSQSKNTWYIDEVAEDPAKSLTEISTDFDRSSPPLQPPPVNSL 403  
Db 351 RSELYTLNGSSVDSQPOSQSKNTWYIDEVAEDPAKSLTEISTDFDRSSPPLQPPPVNSL 410  
QY 404 TTNRFHSLPSLTKMPNTNGSIGHSPLSLSAQSVMBELNAPVQESPLAMPNGSHGL 463  
Db 411 TTNRFHSLPSLTKMPNTNGSIGHSPLSLSAQSVMBELNAPVQESPLAMPNGSHGL 470  
QY 464 EVGSLAEVKNPPFYGVIRWIGQPPGLNEVLAGELEDEACAGCTDGTFRGTRYFTCALKK 523  
Db 471 EVGSLAEVKNPPFYGVIRWIGQPPGLNEVLAGELEDEACAGCTDGTFRGTRYFTCALKK 530  
QY 524 ALFVKLSCRPDSRFASLPQVSNQIERCNSLAFGGYLSVEVEENTPPKMEKEGLEIMIGK 583  
Db 531 ALFVKLSCRPDSRFASLPQVSNQIERCNSLAFGGYLSVEVEENTPPKMEKEGLEIMIGK 590  
QY 584 KKGIOGHVNSCYLDSTLFCFLAFSSVLDVLLRPKEKNDVEYSETQELLRTIENVPLRI 643  
Db 591 KKGIOGHVNSCYLDSTLFCFLAFSSVLDVLLRPKEKNDVEYSETQELLRTIENVPLRI 650  
QY 644 YGYVCATKIMKRLKILEKVEAASGFTSEKDPPEEFLNLFPHILRVEPLLKIRSAGQKVQ 703  
Db 651 YGYVCATKIMKRLKILEKVEAASGFTSEKDPPEEFLNLFPHILRVEPLLKIRSAGQKVQ 710  
QY 704 DCYFQIIFMEKNEKVGVPITIOQLLEWSFINSNLKFAEAPSCLIIOMPRFGKDFLKKIF 763

Db 711 DCYFYQIPMEKNEKVGVTIQOLLEWSFINSNLKFAEAPSCLIIOMPREGDKFLFKKIF 770  
Qy 764 PSLELNITDLEDTTPQCRI CGGLAMYECRECYDDPDISAGIKIQFCKTCNTQVHLHPKR 823  
Db 771 PSLELNITDLEDTTPQCRI CGGLAMYECRECYDDPDISAGIKIQFCKTCNTQVHLHPKR 830  
Qy 824 LNHKYNPVS LPKDLPDMDWRHGCIPCONMELFAVLCIETSHYVAFVKYKGDSDAWLFFDS 883  
Db 831 LNHKYNPVS LPKDLPDMDWRHGCIPCONMELFAVLCIETSHYVAFVKYKGDSDAWLFFDS 890  
Qy 884 MADRDGGQGNIPQVTPCPEVGEY LKMSLEDLHSLDSRRIOGCARRLLCDAYMCYQSP 943  
Db 891 MADRDGGQGNIPQVTPCPEVGEY LKMSLEDLHSLDSRRIOGCARRLLCDAYMCYQSP 950  
Qy 944 TMSLYK 949  
Db 951 TMSLYK 956

RESULT 15  
US-10-755-889-490  
; Sequence 490, Application US/10755889  
; GENERAL INFORMATION:  
; APPLICANT: Bristol-Myers Squibb Company  
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES ASSOCIATED WITH THE NF-KB  
; TITLE OF INVENTION: PATHWAY  
; FILE REFERENCE: D0284 NP  
; CURRENT APPLICATION NUMBER: US/10/755,889  
; PRIOR FILING DATE: 2004-01-13  
; PRIOR APPLICATION NUMBER: U.S. 60/440,068  
; PRIOR FILING DATE: 2003-01-14  
; PRIOR APPLICATION NUMBER: U.S. 60/469,757  
; PRIOR FILING DATE: 2003-05-12  
; NUMBER OF SEQ ID NOS: 823  
; SOFTWARE: Patent in version 3.2  
; SEQ ID NO 490  
; LENGTH: 956  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-755-889-490

Query Match 63.9%; Score 606; DB 33; Length 956;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 606; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 344 RSELYTLNGSSVDSQPQSKNTWYIDEVAEDPAKSLTEISTDFDRSSPPLQPPPVNSL 403  
Db 351 RSELYTLNGSSVDSQPQSKNTWYIDEVAEDPAKSLTEISTDFDRSSPPLQPPPVNSL 410  
Qy 404 TTEHRFSLPFLTKMPNTNGSIGHSPLSLSAQSVMEELNTPAQESPPAMPNGSHGL 463  
Db 411 TTEHRFSLPFLTKMPNTNGSIGHSPLSLSAQSVMEELNTPAQESPPAMPNGSHGL 470  
Qy 464 EVGSLAEVKNPPFFGVIRWIGQPGLAEVLAGLEDEDCAGCTDGTFRGTRYFTCALKK 523  
Db 471 EVGSLAEVKNPPFFGVIRWIGQPGLAEVLAGLEDEDCAGCTDGTFRGTRYFTCALKK 530  
Qy 524 ALFVKLSKCRPDSRFASLPQVSNQIERCNSLAFGGYLSVEVEENTPPKMEKEGLEIMICK 583  
Db 531 ALFVKLSKCRPDSRFASLPQVSNQIERCNSLAFGGYLSVEVEENTPPKMEKEGLEIMICK 590  
Qy 584 KGIQGHYNSCYLDSTLFCFAFSSVLDTVLLRPKEKNDVEYYSQELLRTIENVPLRI 643  
Db 591 KGIQGHYNSCYLDSTLFCFAFSSVLDTVLLRPKEKNDVEYYSQELLRTIENVPLRI 650  
Qy 644 YGYVCATKIMKLRKILEKVEAASGFTSEKDEEFNLILFHHILRVEPLLKIRSAGQKVQ 703  
Db 651 YGYVCATKIMKLRKILEKVEAASGFTSEKDEEFNLILFHHILRVEPLLKIRSAGQKVQ 710  
Qy 704 DCYFYQIPMEKNEKVGVTIQOLLEWSFINSNLKFAEAPSCLIIOMPREGDKFLFKKIF 763  
Db 711 DCYFYQIPMEKNEKVGVTIQOLLEWSFINSNLKFAEAPSCLIIOMPREGDKFLFKKIF 770

Qy 764 PSLELNITDLEDTTPQCRI CGGLAMYECRECYDDPDISAGIKIQFCKTCNTQVHLHPKR 823  
Db 771 PSLELNITDLEDTTPQCRI CGGLAMYECRECYDDPDISAGIKIQFCKTCNTQVHLHPKR 830  
Qy 824 LNHKYNPVS LPKDLPDMDWRHGCIPCONMELFAVLCIETSHYVAFVKYKGDSDAWLFFDS 883  
Db 831 LNHKYNPVS LPKDLPDMDWRHGCIPCONMELFAVLCIETSHYVAFVKYKGDSDAWLFFDS 890  
Qy 884 MADRDGGQGNIPQVTPCPEVGEY LKMSLEDLHSLDSRRIOGCARRLLCDAYMCYQSP 943  
Db 891 MADRDGGQGNIPQVTPCPEVGEY LKMSLEDLHSLDSRRIOGCARRLLCDAYMCYQSP 950  
Qy 944 TMSLYK 949  
Db 951 TMSLYK 956

Search completed: April 18, 2005, 09:57:00  
Job time : 358 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 18, 2005, 09:39:24 ; Search time 22 Seconds  
(without alignments)  
3049.155 Million cell updates/sec

Title: US-09-671-687A-3  
Perfect score: 949  
Sequence: 1 MSSGLWSQEKVTSWPWEERI.....RLLCDAYCMYQSPMSLYK 949

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0  
Searched: 308579 seqs, 70686408 residues

Word size : 0  
Total number of hits satisfying chosen parameters: 308579

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Pending Patents AA New.\*  
1: /cgn2\_6/ptodata/1/paa/PCT NEW COMB.pcp.\*  
2: /cgn2\_6/ptodata/1/paa/US05 NEW COMB.pcp.\*  
3: /cgn2\_6/ptodata/1/paa/US07 NEW COMB.pcp.\*  
4: /cgn2\_6/ptodata/1/paa/US08 NEW COMB.pcp.\*  
5: /cgn2\_6/ptodata/1/paa/US09 NEW COMB.pcp.\*  
6: /cgn2\_6/ptodata/1/paa/US10 NEW COMB.pcp.\*  
7: /cgn2\_6/ptodata/1/paa/US11 NEW COMB.pcp.\*  
8: /cgn2\_6/ptodata/1/paa/US60 NEW COMB.pcp.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	606	63.9	956	US-10-760-678-4	Sequence 4, Appli
2	8	0.8	190	US-60-655-875-14489	Sequence 14489,
3	8	0.8	2216	US-11-031-175-12221	Sequence 12221, A
4	7	0.7	48	US-10-472-963-633	Sequence 633, App
5	7	0.7	107	PCT-US04-26116-116	Sequence 116, App
6	7	0.7	107	US-10-916-840-116	Sequence 116, App
7	7	0.7	112	US-11-031-175-14754	Sequence 14754, A
8	7	0.7	116	PCT-US04-42360-330	Sequence 330, App
9	7	0.7	119	PCT-US05-01983-59	Sequence 59, Appl
10	7	0.7	122	PCT-US04-09510-1133	Sequence 1133, Ap
11	7	0.7	123	PCT-US04-09510-1134	Sequence 1134, Ap
12	7	0.7	125	US-11-027-399-4643	Sequence 4643, Ap
13	7	0.7	125	US-11-027-843-4643	Sequence 4643, Ap
14	7	0.7	125	US-11-027-878-4643	Sequence 4643, Ap
15	7	0.7	125	US-11-028-169-4643	Sequence 4643, Ap
16	7	0.7	125	US-11-028-204-4643	Sequence 4643, Ap
17	7	0.7	125	US-11-027-877-4643	Sequence 4643, Ap
18	7	0.7	125	US-11-027-879-4643	Sequence 4643, Ap
19	7	0.7	125	US-11-028-149-4643	Sequence 4643, Ap
20	7	0.7	125	US-11-027-802-4643	Sequence 4643, Ap
21	7	0.7	125	US-11-027-890-4643	Sequence 4643, Ap
22	7	0.7	125	US-11-027-892-4643	Sequence 4643, Ap
23	7	0.7	125	US-11-028-099-4643	Sequence 4643, Ap
24	7	0.7	125	US-11-028-197-4643	Sequence 4643, Ap
25	7	0.7	125	US-11-027-844-4643	Sequence 4643, Ap

ALIGNMENTS

RESULT 1  
US-10-760-678-4  
; Sequence 4, Application US/10760678  
; GENERAL INFORMATION:  
; APPLICANT: Derry, Jonathan  
; APPLICANT: Fanslow, William  
; APPLICANT: Dougall, William  
; TITLE OF INVENTION: SCREENING ASSAYS FOR AGONISTS OR ANTAGONISTS OF CD40 SIGNALING  
; FILE REFERENCE: 3198  
; CURRENT APPLICATION NUMBER: US/10/760,678  
; CURRENT FILING DATE: 2004-01-20  
; PRIOR APPLICATION NUMBER: US/09/851,673  
; PRIOR FILING DATE: 2001-05-08  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: Patent in version 3.0  
; SEQ ID NO 4  
; LENGTH: 956  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-760-678-4

Query Match 63.9%; Score 606; DB 6; Length 956;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 606; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	344	RSELFYTLNGSSVDSQPQSKSNTWTYDEVAEDPAKSLTEISTDFDRSSPPLQPPVNSL	403
Db	351	RSELFYTLNGSSVDSQPQSKSNTWTYDEVAEDPAKSLTEISTDFDRSSPPLQPPVNSL	410
QY	404	TTENRPHSLPFSLTCKMPTNGSIGHSPLSLSAQSVMEELNTAPVQSPPLAMPGGNSHGL	463
Db	411	TTENRPHSLPFSLTCKMPTNGSIGHSPLSLSAQSVMEELNTAPVQSPPLAMPGGNSHGL	470
QY	464	EVGSLAEVKNENPPFYGVIRWIGOPGLNEVLAGELEDECACTDGTFRGTRYFTCALKK	523
Db	471	EVGSLAEVKNENPPFYGVIRWIGOPGLNEVLAGELEDECACTDGTFRGTRYFTCALKK	530
QY	524	ALFVKLKS CRPSRFPASLPVSNQIERCNSLAPGGYLSVSEVENTPPKMEKEGLEIMIGK	583
Db	531	ALFVKLKS CRPSRFPASLPVSNQIERCNSLAPGGYLSVSEVENTPPKMEKEGLEIMIGK	590
QY	584	KGIGOHYNSCYLDSLTFCLFAFSSVLDTVLLRPKKNDEYVYSETOELLRTIIVNPLRI	643
Db	591	KGIGOHYNSCYLDSLTFCLFAFSSVLDTVLLRPKKNDEYVYSETOELLRTIIVNPLRI	650
QY	644	YGVCATKIMKLKRLKLEKVEAASGFTSEBKDPPEEFNILPHILRVEPLLIKRSAGQKVQ	703
Db	651	YGVCATKIMKLKRLKLEKVEAASGFTSEBKDPPEEFNILPHILRVEPLLIKRSAGQKVQ	710

QY 704 DCYFQIFMEKNEKVGVTIQQLEWSFINSNLKFAEAPSCLIQMPRFGKDFKFKKIF 763  
DB 711 DCYFQIFMEKNEKVGVTIQQLEWSFINSNLKFAEAPSCLIQMPRFGKDFKFKKIF 770  
QY 764 PSLEINTDLEDTPRQCRI CGGLAMYECRECYDDPDISAGKIKQFCCTCNTQVHLHPR 823  
DB 771 PSLEINTDLEDTPRQCRI CGGLAMYECRECYDDPDISAGKIKQFCCTCNTQVHLHPR 830  
QY 824 LNHKYNPVSLEKPLDPDWRHGCIPQNNELFAVLCIETSHYVAVKYGKDDSAWLPFDS 883  
DB 831 LNHKYNPVSLEKPLDPDWRHGCIPQNNELFAVLCIETSHYVAVKYGKDDSAWLPFDS 890  
QY 884 MADRDGGQGNIPQVTFPCPEVGEYKMSLEDLHSLDRRIQGCARRLLCDAYMCWQSP 943  
DB 891 MADRDGGQGNIPQVTFPCPEVGEYKMSLEDLHSLDRRIQGCARRLLCDAYMCWQSP 950  
QY 944 TMSLYK 949  
DB 951 TMSLYK 956

## RESULT 2

US-60-655-875-144489  
; Sequence 144489, Application US/60655875  
; GENERAL INFORMATION:  
; APPLICANT: Boukharov, Andrey  
; APPLICANT: Du, Zijing  
; APPLICANT: Guo, Liang  
; APPLICANT: Kovalic, David  
; APPLICANT: Lu, Maolong  
; APPLICANT: McCarter, James  
; APPLICANT: Miller, Nancy  
; APPLICANT: Williams, Deryck  
; APPLICANT: Vaudin, Mark  
; APPLICANT: Wu, Wei  
; TITLE OF INVENTION: METHODS FOR GENETIC CONTROL OF HETERODERA INFESTATIONS  
; TITLE OF INVENTION: IN PLANTS AND COMPOSITIONS THEREOF  
; FILE REFERENCE: 38-21(53885)  
; CURRENT APPLICATION NUMBER: US/60/655,875  
; CURRENT FILING DATE: 2005-02-24  
; NUMBER OF SEQ ID NOS: 171306  
; SEQ ID NO 144489  
; LENGTH: 190  
; TYPE: PRT  
; ORGANISM: Heterodera glycines  
; FEATURE:  
; OTHER INFORMATION: Coding regions on vcdna: vcdna=SeqID\_70912; Strand=-; Position=1  
US-60-655-875-144489

Query Match 0.8%; Score 8; DB 8; Length 190;  
Best Local Similarity 100.0%; Pred. No. 9.9;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 617 PKEKNDVE 624  
DB 173 PKEKNDVE 180

## RESULT 3

US-11-031-175-12221  
; Sequence 12221, Application US/11031175  
; GENERAL INFORMATION:  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Wiegand, Roger C.  
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof  
; FILE REFERENCE: 38-10(15849)B  
; CURRENT APPLICATION NUMBER: US/11/031,175  
; CURRENT FILING DATE: 2005-01-08  
; PRIOR APPLICATION NUMBER: 60/217,883

; PRIOR FILING DATE: 2000-07-10  
; NUMBER OF SEQ ID NOS: 16825  
; SEQ ID NO 12221  
; LENGTH: 2216  
; TYPE: PRT  
; ORGANISM: Myxococcus xanthus  
US-11-031-175-12221

Query Match 0.8%; Score 8; DB 7; Length 2216;  
Best Local Similarity 100.0%; Pred. No. 99;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 169 EGRGQGT 176  
DB 395 EGRGQGT 402

## RESULT 4

US-10-472-963-633  
; Sequence 633, Application US/10472963  
; GENERAL INFORMATION:  
; APPLICANT: Human Genome Sciences, Inc.  
; TITLE OF INVENTION: Human Secreted Proteins  
; FILE REFERENCE: PS954PCT  
; CURRENT APPLICATION NUMBER: US/10/472,963  
; CURRENT FILING DATE: 2003-09-25  
; PRIOR APPLICATION NUMBER: PCT/US02/09370  
; PRIOR FILING DATE: 2002-03-26  
; PRIOR APPLICATION NUMBER: US 60/278,650  
; PRIOR FILING DATE: 2001-03-27  
; PRIOR APPLICATION NUMBER: US 09/950,082  
; PRIOR FILING DATE: 2001-09-12  
; PRIOR APPLICATION NUMBER: US 09/950,083  
; PRIOR FILING DATE: 2001-09-12  
; NUMBER OF SEQ ID NOS: 1834  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 633  
; LENGTH: 48  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-472-963-633

Query Match 0.7%; Score 7; DB 6; Length 48;  
Best Local Similarity 100.0%; Pred. No. 30;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 580 MIGKKKG 586  
DB 1 MIGKKKG 7

## RESULT 5

PCT-US04-26116-116  
; Sequence 116, Application PC/TUS0426116  
; GENERAL INFORMATION:  
; APPLICANT: Dyax Corp.  
; TITLE OF INVENTION: TIEI-BINDING LIGANDS  
; FILE REFERENCE: 10280-083wo11  
; CURRENT APPLICATION NUMBER: PCT/US04/26116  
; CURRENT FILING DATE: 2004-08-12  
; PRIOR APPLICATION NUMBER: US 60/494,713  
; PRIOR FILING DATE: 2003-08-12  
; NUMBER OF SEQ ID NOS: 161  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 116  
; LENGTH: 107  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Synthetically generated peptide  
PCT-US04-26116-116

Query Match 0.7%; Score 7; DB 1; Length 107;

Best Local Similarity 100.0%; Pred. No. 63;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 429 SPLSLSA 435  
Db 7 SPLSLSA 13

RESULT 6  
US-10-916-840-116  
; Sequence 116, Application US/10916840  
; GENERAL INFORMATION:  
; APPLICANT: Hufton, Simon E.  
; APPLICANT: Hoet, Rene  
; APPLICANT: Pieters, Henk  
; APPLICANT: Kent, Rachel Baribault  
; APPLICANT: Rookey, Kristen  
; TITLE OF INVENTION: TIE1-BINDING LIGANDS  
; FILE REFERENCE: 10280-083001  
; CURRENT APPLICATION NUMBER: US/10/916,840  
; PRIOR FILING DATE: 2004-08-12  
; PRIOR FILING DATE: 2004-08-12  
; PRIOR FILING DATE: 2003-08-12  
; NUMBER OF SEQ ID NOS: 161  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 116  
; LENGTH: 107  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Synthetically generated peptide  
US-10-916-840-116

Query Match 0.7%; Score 7; DB 6; Length 107;  
Best Local Similarity 100.0%; Pred. No. 63;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 429 SPLSLSA 435  
Db 7 SPLSLSA 13

RESULT 7  
US-11-031-175-14754  
; Sequence 14754, Application US/11031175  
; GENERAL INFORMATION:  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Wiegand, Roger C.  
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof  
; FILE REFERENCE: 38-10(15849)B  
; CURRENT APPLICATION NUMBER: US/11/031,175  
; CURRENT FILING DATE: 2005-01-08  
; PRIOR FILING DATE: 2005-01-08  
; PRIOR FILING DATE: 2000-07-10  
; NUMBER OF SEQ ID NOS: 16825  
; SEQ ID NO 14754  
; LENGTH: 112  
; TYPE: PRT  
; ORGANISM: Myxococcus xanthus  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(112)  
; OTHER INFORMATION: unsure at all Xaa locations  
US-11-031-175-14754

Query Match 0.7%; Score 7; DB 7; Length 112;  
Best Local Similarity 100.0%; Pred. No. 66;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 696 RSAGQKV 702  
Db 7 RSAGQKV 13

Db 73 RSAGQKV 79

RESULT 8  
PCT-US04-42360-330  
; Sequence 330, Application PC/TUS0442360  
; GENERAL INFORMATION:  
; APPLICANT: The General Hospital Corporation doing business as Massachusetts General  
; APPLICANT: Hospital / Bayer AG  
; TITLE OF INVENTION: Nucleotide sequences involved in pain  
; FILE REFERENCE: 17633/2048  
; CURRENT APPLICATION NUMBER: PCT/US04/42360  
; CURRENT FILING DATE: 2004-12-17  
; PRIOR APPLICATION NUMBER: 60/531,341  
; PRIOR FILING DATE: 19-Dec-2003  
; NUMBER OF SEQ ID NOS: 2587  
; SOFTWARE: Perl script  
; SEQ ID NO 330  
; LENGTH: 116  
; TYPE: PRT  
; ORGANISM: Rattus norvegicus  
; PUBLICATION INFORMATION:  
; DATABASE ACCESSION NUMBER: Refseq / NP\_062035  
; DATABASE ENTRY DATE: 2003-10-06  
PCT-US04-42360-330

Query Match 0.7%; Score 7; DB 1; Length 116;  
Best Local Similarity 100.0%; Pred. No. 68;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 563 VVEENTP 569  
Db 104 VVEENTP 110

RESULT 9  
PCT-US05-01983-59  
; Sequence 59, Application PC/TUS0501983  
; GENERAL INFORMATION:  
; APPLICANT: Solazyme, Inc.  
; APPLICANT: Dillon, Harrison F.  
; TITLE OF INVENTION: Methods and Compositions for Evolving Microbial Hydrogen  
; TITLE OF INVENTION: Production  
; FILE REFERENCE: H2042101-CIP  
; CURRENT APPLICATION NUMBER: PCT/US05/01983  
; CURRENT FILING DATE: 2005-01-31  
; PRIOR APPLICATION NUMBER: US 10/287,750  
; PRIOR FILING DATE: 2002-11-04  
; PRIOR APPLICATION NUMBER: US 10/411,910  
; PRIOR FILING DATE: 2003-04-12  
; PRIOR APPLICATION NUMBER: US 60/500,032  
; PRIOR FILING DATE: 2003-09-03  
; NUMBER OF SEQ ID NOS: 184  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 59  
; LENGTH: 119  
; TYPE: PRT  
; ORGANISM: Metopus contortus  
PCT-US05-01983-59

Query Match 0.7%; Score 7; DB 1; Length 119;  
Best Local Similarity 100.0%; Pred. No. 70;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 829 NPVSLPK 835  
Db 93 NPVSLPK 99

RESULT 10  
PCT-US04-09510-1133  
; Sequence 1133, Application PC/TUS0409510  
; GENERAL INFORMATION:

; APPLICANT: Epimmune Inc.  
 ; APPLICANT: Baker, Denise M.  
 ; APPLICANT: Livingston, Brian D.  
 ; APPLICANT: Chesnut, Robert W.  
 ; APPLICANT: Sette, Alessandro  
 ; APPLICANT: Newman, Mark J.  
 ; TITLE OF INVENTION: Methods of Identifying Optimal Variants of Peptide Epitopes  
 ; FILE REFERENCE: 2060.026PC01  
 ; CURRENT APPLICATION NUMBER: PCT/US04/09510  
 ; CURRENT FILING DATE: 2004-03-29  
 ; PRIOR APPLICATION NUMBER: 60/458,026  
 ; PRIOR FILING DATE: 2003-03-28  
 ; NUMBER OF SEQ ID NOS: 1946  
 ; SOFTWARE: PatentIn version 3.3  
 ; SEQ ID NO 1133  
 ; LENGTH: 122  
 ; TYPE: PRT  
 ; ORGANISM: Artificial  
 ; FEATURE:  
 ; OTHER INFORMATION: HIV Rev  
 ; NAME/KEY: misc feature  
 ; LOCATION: (123)-(122)  
 ; OTHER INFORMATION: Xaa can be any naturally occurring amino acid  
 PCT-US04-09510-1133

Query Match 0.7%; Score 7; DB 1; Length 122;  
 Best Local Similarity 100.0%; Pred. No. 71;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 223 VELPPL 229  
 Db 72 VELPPL 78

RESULT 11  
 PCT-US04-09510-1134  
 ; Sequence 1134, Application PC/TUS0409510  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Epimmune Inc.  
 ; APPLICANT: Baker, Denise M.  
 ; APPLICANT: Livingston, Brian D.  
 ; APPLICANT: Chesnut, Robert W.  
 ; APPLICANT: Sette, Alessandro  
 ; APPLICANT: Newman, Mark J.  
 ; TITLE OF INVENTION: Methods of Identifying Optimal Variants of Peptide Epitopes  
 ; FILE REFERENCE: 2060.026PC01  
 ; CURRENT APPLICATION NUMBER: PCT/US04/09510  
 ; CURRENT FILING DATE: 2004-03-29  
 ; PRIOR APPLICATION NUMBER: 60/458,026  
 ; PRIOR FILING DATE: 2003-03-28  
 ; NUMBER OF SEQ ID NOS: 1946  
 ; SOFTWARE: PatentIn version 3.3  
 ; SEQ ID NO 1134  
 ; LENGTH: 123  
 ; TYPE: PRT  
 ; ORGANISM: Artificial  
 ; FEATURE:  
 ; OTHER INFORMATION: HIV Rev  
 ; NAME/KEY: misc feature  
 ; LOCATION: (123)-(123)  
 ; OTHER INFORMATION: Xaa can be any naturally occurring amino acid  
 PCT-US04-09510-1134

Query Match 0.7%; Score 7; DB 1; Length 123;  
 Best Local Similarity 100.0%; Pred. No. 72;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 223 VELPPL 229  
 Db 73 VELPPL 79

RESULT 12  
 US-11-027-399-4643  
 ; Sequence 4643, Application US/11027399  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Doucette-Stamm, Lynn  
 ; APPLICANT: Bush, David  
 ; APPLICANT: Zeng, Qiangdong  
 ; APPLICANT: Opperman, Timothy  
 ; APPLICANT: Houseweart, Chad Eric  
 ; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus  
 ; FILE REFERENCE: 3687.1000-015  
 ; CURRENT APPLICATION NUMBER: US/11/027,399  
 ; CURRENT FILING DATE: 2004-12-30  
 ; PRIOR APPLICATION NUMBER: US 10/640,833  
 ; PRIOR FILING DATE: 2003-08-14  
 ; PRIOR APPLICATION NUMBER: US 09/583,110  
 ; PRIOR FILING DATE: 2000-05-26  
 ; PRIOR APPLICATION NUMBER: US 09/107,433  
 ; PRIOR FILING DATE: 1998-06-30  
 ; PRIOR APPLICATION NUMBER: US 60/085,131  
 ; PRIOR FILING DATE: 1998-05-12  
 ; PRIOR APPLICATION NUMBER: US 60/051,553  
 ; PRIOR FILING DATE: 1997-07-02  
 ; NUMBER OF SEQ ID NOS: 5322  
 ; SEQ ID NO 4643  
 ; LENGTH: 125  
 ; TYPE: PRT  
 ; ORGANISM: Streptococcus pneumoniae  
 US-11-027-399-4643

Query Match 0.7%; Score 7; DB 7; Length 125;  
 Best Local Similarity 100.0%; Pred. No. 73;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 688 RVEPLLK 694  
 Db 112 RVEPLLK 118

RESULT 13  
 US-11-027-843-4643  
 ; Sequence 4643, Application US/11027843  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Doucette-Stamm, Lynn  
 ; APPLICANT: Bush, David  
 ; APPLICANT: Zeng, Qiangdong  
 ; APPLICANT: Opperman, Timothy  
 ; APPLICANT: Houseweart, Chad Eric  
 ; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus  
 ; FILE REFERENCE: 3687.1000-024  
 ; CURRENT APPLICATION NUMBER: US/11/027,843  
 ; CURRENT FILING DATE: 2004-12-30  
 ; PRIOR APPLICATION NUMBER: US 10/640,833  
 ; PRIOR FILING DATE: 2003-08-14  
 ; PRIOR APPLICATION NUMBER: US 09/583,110  
 ; PRIOR FILING DATE: 2000-05-26  
 ; PRIOR APPLICATION NUMBER: US 09/107,433  
 ; PRIOR FILING DATE: 1998-06-30  
 ; PRIOR APPLICATION NUMBER: US 60/085,131  
 ; PRIOR FILING DATE: 1998-05-12  
 ; PRIOR APPLICATION NUMBER: US 60/051,553  
 ; PRIOR FILING DATE: 1997-07-02  
 ; NUMBER OF SEQ ID NOS: 5322  
 ; SEQ ID NO 4643  
 ; LENGTH: 125  
 ; TYPE: PRT  
 ; ORGANISM: Streptococcus pneumoniae  
 US-11-027-843-4643

Query Match 0.7%; Score 7; DB 7; Length 125;

```
Best Local Similarity 100.0%; Pred. No. 73;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      688 RVEPLLK 694
Db      112 RVEPLLK 118

RESULT 14
US-11-027-878-4643
; Sequence 4643, Application US/11027878
; GENERAL INFORMATION:
; APPLICANT: Doucette-Stamm, Lynn
; APPLICANT: Bush, David
; APPLICANT: Zeng, Qiaodong
; APPLICANT: Opperman, Timothy
; APPLICANT: Houseweart, Chad Eric
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
; FILE REFERENCE: 3687.1000-008
; CURRENT APPLICATION NUMBER: US/11/027,878
; CURRENT FILING DATE: 2004-12-30
; PRIOR APPLICATION NUMBER: US 10/640,833
; PRIOR FILING DATE: 2003-08-14
; PRIOR APPLICATION NUMBER: US 09/583,110
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/107,433
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/085,131
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: US 60/051,553
; PRIOR FILING DATE: 1997-07-02
; NUMBER OF SEQ ID NOS: 5322
; SEQ ID NO 4643
; LENGTH: 125
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-11-027-878-4643

Query Match      0.7%; Score 7; DB 7; Length 125;
Best Local Similarity 100.0%; Pred. No. 73;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      688 RVEPLLK 694
Db      112 RVEPLLK 118

RESULT 15
US-11-028-169-4643
; Sequence 4643, Application US/11028169
; GENERAL INFORMATION:
; APPLICANT: Doucette-Stamm, Lynn
; APPLICANT: Bush, David
; APPLICANT: Zeng, Qiaodong
; APPLICANT: Opperman, Timothy
; APPLICANT: Houseweart, Chad Eric
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
; FILE REFERENCE: 3687.1000-010
; CURRENT APPLICATION NUMBER: US/11/028,169
; CURRENT FILING DATE: 2004-12-30
; PRIOR APPLICATION NUMBER: US 10/640,833
; PRIOR FILING DATE: 2003-08-14
; PRIOR APPLICATION NUMBER: US 09/583,110
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/107,433
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/085,131
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: US 60/051,553
; PRIOR FILING DATE: 1997-07-02
; NUMBER OF SEQ ID NOS: 5322
```

```
; SEQ ID NO 4643
; LENGTH: 125
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-11-028-169-4643

Query Match      0.7%; Score 7; DB 7; Length 125;
Best Local Similarity 100.0%; Pred. No. 73;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      688 RVEPLLK 694
Db      112 RVEPLLK 118

Search completed: April 18, 2005, 09:57:31
Job time : 23 secs
```



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 18, 2005, 09:40:40 ; Search time 57 Seconds  
(without alignments)  
5533.680 Million cell updates/sec

Title: US-09-671-687A-3  
Perfect score: 949  
Sequence: 1 MSSGLWSQKVTSPWEERI.....RLLCDAYMCVQPTMSLYK 949

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 1421835 seqs, 332370683 residues

Word size : 0  
Total number of hits satisfying chosen parameters: 1421835

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

- Database : Published Applications AA.\*
- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
  - 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*
  - 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
  - 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*
  - 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
  - 6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*
  - 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
  - 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
  - 9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*
  - 10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*
  - 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*
  - 12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
  - 13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*
  - 14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*
  - 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*
  - 16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pep.\*
  - 17: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*
  - 18: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep.\*
  - 19: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*
  - 20: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Query %	ID	Description
1	606	63.9	731	16	US-10-921-707-9
2	606	63.9	956	10	US-09-851-673-4
3	606	63.9	956	16	US-10-755-889-490
4	606	63.9	956	16	US-10-788-792-250
5	208	21.9	261	15	US-10-264-237-1609
6	148	15.6	238	15	US-10-264-237-1610
7	113	11.9	113	9	US-09-864-761-34675
8	8	0.8	49	15	US-10-424-599-186263
9	8	0.8	78	14	US-10-058-053A-218
10	8	0.8	78	16	US-10-838-226-218
11	8	0.8	122	16	US-10-437-963-180681
12	8	0.8	143	15	US-10-424-599-170638
13	8	0.8	188	15	US-10-424-599-159539

14	8	0.8	371	16	US-10-437-963-188186	Sequence 188186,
15	8	0.8	450	15	US-10-282-122A-60805	Sequence 60805, A
16	8	0.8	1086	15	US-10-369-493-5963	Sequence 5963, Ap
17	7	0.7	28	10	US-09-983-802-607	Sequence 607, App
18	7	0.7	28	10	US-09-984-490-607	Sequence 607, App
19	7	0.7	28	11	US-09-973-278-461	Sequence 461, App
20	7	0.7	30	16	US-10-632-706-150	Sequence 150, App
21	7	0.7	30	16	US-10-632-706-154	Sequence 154, App
22	7	0.7	33	14	US-10-081-872-219	Sequence 219, App
23	7	0.7	33	15	US-10-385-305-219	Sequence 219, App
24	7	0.7	45	15	US-10-264-049-3820	Sequence 3820, Ap
25	7	0.7	53	15	US-10-613-413A-73	Sequence 73, Appl
26	7	0.7	53	17	US-10-885-225-73	Sequence 73, Appl
27	7	0.7	61	15	US-10-424-599-276896	Sequence 276896,
28	7	0.7	62	15	US-10-424-599-193894	Sequence 193894,
29	7	0.7	62	15	US-10-424-599-221399	Sequence 221399,
30	7	0.7	67	11	US-09-864-408A-6310	Sequence 6310, Ap
31	7	0.7	67	15	US-10-424-599-266715	Sequence 266715,
32	7	0.7	73	16	US-10-767-701-49557	Sequence 49557, A
33	7	0.7	79	16	US-10-437-963-184451	Sequence 184451,
34	7	0.7	82	16	US-10-767-701-61504	Sequence 61504, A
35	7	0.7	86	15	US-10-424-599-150683	Sequence 150683,
36	7	0.7	86	16	US-10-437-963-195228	Sequence 195228,
37	7	0.7	88	16	US-10-437-963-107983	Sequence 107983,
38	7	0.7	92	16	US-10-437-963-162456	Sequence 162456,
39	7	0.7	98	15	US-10-424-599-210307	Sequence 210307,
40	7	0.7	102	16	US-10-437-963-106579	Sequence 106579,
41	7	0.7	104	15	US-10-424-599-247009	Sequence 247009,
42	7	0.7	104	16	US-10-767-701-48578	Sequence 48578, A
43	7	0.7	106	11	US-09-864-408A-2374	Sequence 2374, Ap
44	7	0.7	106	16	US-10-437-963-199204	Sequence 199204,
45	7	0.7	107	16	US-10-663-244-83	Sequence 83, Appl

ALIGNMENTS

RESULT 1  
US-10-921-707-9  
; Sequence 9, Application US/10921707  
; Publication No. US20050003447A1  
; GENERAL INFORMATION:  
; APPLICANT: INCYTE PHARMACEUTICALS, INC.  
; APPLICANT: LAL, Preeti  
; APPLICANT: TANG, Y. Tom  
; APPLICANT: YUE, Henry  
; APPLICANT: HILLMAN, Jennifer L.  
; APPLICANT: BANDMAN, Olga  
; APPLICANT: CORLEY, Neil C.  
; APPLICANT: GUEGLER, Karl J.  
; APPLICANT: PATTERSON, Chandra  
; APPLICANT: AZIMZAI, Yalda  
; APPLICANT: BAUGHN, Mariah R.  
; TITLE OF INVENTION: HUMAN CYTOSKELETON ASSOCIATED PROTEINS  
; FILE REFERENCE: PF-0594 PCT  
; CURRENT APPLICATION NUMBER: US/10/921,707  
; CURRENT FILING DATE: 2004-08-19  
; PRIOR APPLICATION NUMBER: US/09/786,797  
; PRIOR FILING DATE: 2001-06-25  
; PRIOR APPLICATION NUMBER: 09/156,470; unassigned; 60/131,321  
; PRIOR FILING DATE: 1998-09-18; 1998-09-18; 1999-04-27  
; NUMBER OF SEQ ID NOS: 32  
; SOFTWARE: PERL Program  
; SEQ ID NO 9  
; LENGTH: 731  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; OTHER INFORMATION: Incyte ID No: 2363327  
US-10-921-707-9  
Query Match 63.9%; Score 606; DB 16; Length 731;



Db 351 RSELYTLNGSSVDSQOSKSNWYIDEVAEDPAKSLTEISTDFDRSSPPIQPPVNSL 410  
 Qy 404 TTRNRFHSLPFSLTAKMPNTNGSIHSPLSLSAQSVMEELNAPVQESPLAMPNGSHGL 463  
 Db 411 TTRNRFHSLPFSLTAKMPNTNGSIHSPLSLSAQSVMEELNAPVQESPLAMPNGSHGL 470  
 Qy 464 EVGSLAEVKNPPFFGVIRWIGOPGLNEVLAGLEDEBACAGCTDGTFRGTRYFTCALKK 523  
 Db 471 EVGSLAEVKNPPFFGVIRWIGOPGLNEVLAGLEDEBACAGCTDGTFRGTRYFTCALKK 530  
 Qy 524 ALFVKLKSRCRPSDRFASLQPVSNQIERCNSLAFGGVLSVVEENTPPKMEKEGLEIMIGK 583  
 Db 531 ALFVKLKSRCRPSDRFASLQPVSNQIERCNSLAFGGVLSVVEENTPPKMEKEGLEIMIGK 590  
 Qy 584 KKGIOGHYNSCYLDSTLFCFAFSSVLDVTLRPEKKNDEVEYSETQELLRTTEIVNPLRI 643  
 Db 591 KKGIOGHYNSCYLDSTLFCFAFSSVLDVTLRPEKKNDEVEYSETQELLRTTEIVNPLRI 650  
 Qy 644 YGYVCATKIMKRLKILEKVEAASGTSBEKOPPEEFNLILFHHILRVEPLLLKIRSAGQKVQ 703  
 Db 651 YGYVCATKIMKRLKILEKVEAASGTSBEKOPPEEFNLILFHHILRVEPLLLKIRSAGQKVQ 710  
 Qy 704 DCYFYQIPEKNEKVGVPPTIQOLLEWSFINSNLKFAEAPSCILIIQMPRFGKDFLKKIF 763  
 Db 711 DCYFYQIPEKNEKVGVPPTIQOLLEWSFINSNLKFAEAPSCILIIQMPRFGKDFLKKIF 770  
 Qy 764 PSLELNITDLEDDTPRQCRICGGLAMYECCYDDPDISAGKIKQFCCTCNTQVHLHPKR 823  
 Db 771 PSLELNITDLEDDTPRQCRICGGLAMYECCYDDPDISAGKIKQFCCTCNTQVHLHPKR 830  
 Qy 824 LNHNKYNVSLPKDLPDWDRHGCIPQNNMELFAVLCIETSHYVAFVKYKGDSDSAMLFFDS 883  
 Db 831 LNHNKYNVSLPKDLPDWDRHGCIPQNNMELFAVLCIETSHYVAFVKYKGDSDSAMLFFDS 890  
 Qy 884 MADRDGGONGNIPQVTPCPEVGEYVKMSLEDLHSLDSRRIOGCARRLLCDAYMCWQSP 943  
 Db 891 MADRDGGONGNIPQVTPCPEVGEYVKMSLEDLHSLDSRRIOGCARRLLCDAYMCWQSP 950  
 Qy 944 TMSLYK 949  
 Db 951 TMSLYK 956

RESULT 4  
 US-10-788-792-250  
 ; Sequence 250, Application US/10788792  
 ; Publication No. US20040191819A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Bayer Pharmaceuticals Corporation  
 ; APPLICANT: Eveleigh, Deepa  
 ; APPLICANT: Bigwood, Douglas  
 ; TITLE OF INVENTION: EXPRESSION PROFILES FOR BREAST CANCER AND METHODS OF USE  
 ; FILE REFERENCE: 5152  
 ; CURRENT APPLICATION NUMBER: US/10/788,792  
 ; PRIOR FILING DATE: 2004-02-27  
 ; PRIOR APPLICATION NUMBER: US 60/450,655  
 ; PRIOR FILING DATE: 2003-02-28  
 ; NUMBER OF SEQ ID NOS: 254  
 ; SOFTWARE: Patent in version 3.2  
 ; SEQ ID NO 250  
 ; LENGTH: 960  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-10-788-792-250

Query Match 63.9%; Score 606; DB 16; Length 960;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 606; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 344 RSELYTLNGSSVDSQOSKSNWYIDEVAEDPAKSLTEISTDFDRSSPPIQPPVNSL 403  
 Db 355 RSELYTLNGSSVDSQOSKSNWYIDEVAEDPAKSLTEISTDFDRSSPPIQPPVNSL 414

Qy 404 TTRNRFHSLPFSLTAKMPNTNGSIHSPLSLSAQSVMEELNAPVQESPLAMPNGSHGL 463  
 Db 415 TTRNRFHSLPFSLTAKMPNTNGSIHSPLSLSAQSVMEELNAPVQESPLAMPNGSHGL 474  
 Qy 464 EVGSLAEVKNPPFFGVIRWIGOPGLNEVLAGLEDEBACAGCTDGTFRGTRYFTCALKK 523  
 Db 475 EVGSLAEVKNPPFFGVIRWIGOPGLNEVLAGLEDEBACAGCTDGTFRGTRYFTCALKK 534  
 Qy 524 ALFVKLKSRCRPSDRFASLQPVSNQIERCNSLAFGGVLSVVEENTPPKMEKEGLEIMIGK 583  
 Db 535 ALFVKLKSRCRPSDRFASLQPVSNQIERCNSLAFGGVLSVVEENTPPKMEKEGLEIMIGK 594  
 Qy 584 KKGIOGHYNSCYLDSTLFCFAFSSVLDVTLRPEKKNDEVEYSETQELLRTTEIVNPLRI 643  
 Db 595 KKGIOGHYNSCYLDSTLFCFAFSSVLDVTLRPEKKNDEVEYSETQELLRTTEIVNPLRI 654  
 Qy 644 YGYVCATKIMKRLKILEKVEAASGTSBEKOPPEEFNLILFHHILRVEPLLLKIRSAGQKVQ 703  
 Db 655 YGYVCATKIMKRLKILEKVEAASGTSBEKOPPEEFNLILFHHILRVEPLLLKIRSAGQKVQ 714  
 Qy 704 DCYFYQIPEKNEKVGVPPTIQOLLEWSFINSNLKFAEAPSCILIIQMPRFGKDFLKKIF 763  
 Db 715 DCYFYQIPEKNEKVGVPPTIQOLLEWSFINSNLKFAEAPSCILIIQMPRFGKDFLKKIF 774  
 Qy 764 PSLELNITDLEDDTPRQCRICGGLAMYECCYDDPDISAGKIKQFCCTCNTQVHLHPKR 823  
 Db 775 PSLELNITDLEDDTPRQCRICGGLAMYECCYDDPDISAGKIKQFCCTCNTQVHLHPKR 834  
 Qy 824 LNHNKYNVSLPKDLPDWDRHGCIPQNNMELFAVLCIETSHYVAFVKYKGDSDSAMLFFDS 883  
 Db 835 LNHNKYNVSLPKDLPDWDRHGCIPQNNMELFAVLCIETSHYVAFVKYKGDSDSAMLFFDS 894  
 Qy 884 MADRDGGONGNIPQVTPCPEVGEYVKMSLEDLHSLDSRRIOGCARRLLCDAYMCWQSP 943  
 Db 895 MADRDGGONGNIPQVTPCPEVGEYVKMSLEDLHSLDSRRIOGCARRLLCDAYMCWQSP 954  
 Qy 944 TMSLYK 949  
 Db 955 TMSLYK 960

RESULT 5  
 US-10-264-237-1609  
 ; Sequence 1609, Application US/10264237  
 ; Publication No. US20040009491A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Birse et al.  
 ; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
 ; FILE REFERENCE: PA131P1  
 ; CURRENT APPLICATION NUMBER: US/10/264,237  
 ; CURRENT FILING DATE: 2002-10-04  
 ; PRIOR APPLICATION NUMBER: PCT/US01/16450  
 ; PRIOR FILING DATE: 2001-05-18  
 ; PRIOR APPLICATION NUMBER: US 60/205,515  
 ; PRIOR FILING DATE: 2000-05-19  
 ; NUMBER OF SEQ ID NOS: 2876  
 ; SOFTWARE: Patent in Ver. 3.1  
 ; SEQ ID NO 1609  
 ; LENGTH: 261  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: MISC FEATURE  
 ; LOCATION: (209)  
 ; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids  
 ; FEATURE:  
 ; NAME/KEY: MISC FEATURE  
 ; LOCATION: (218)  
 ; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids  
 ; FEATURE:  
 ; NAME/KEY: MISC FEATURE  
 ; LOCATION: (227)  
 ; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids

FEATURE: ; NAME/KEY: MISC FEATURE ; LOCATION: (257) ; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids

US-10-264-237-1609

Query Match 21.9%; Score 208; DB 15; Length 261; Best Local Similarity 100.0%; Pred. No. 1.5e-196; Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 653 MKLRKILKVEAASGFTSEEDPEFLNLFPHILRVLPFLKIRSGAKVQDCYFQIFM 712

Db 1 MKLRKILKVEAASGFTSEEDPEFLNLFPHILRVLPFLKIRSGAKVQDCYFQIFM 60

Qy 713 EKNEKVGVTTIQQLLEWSEFINSNLKFAEAPSCLIIQMPFRGKDFKLFKFIKPSLELNITD 772

Db 61 EKNEKVGVTTIQQLLEWSEFINSNLKFAEAPSCLIIQMPFRGKDFKLFKFIKPSLELNITD 120

Qy 773 LLEDTPROCRICGGLAMVECRECYDDPDISAGIKQFCKTQVHLHPKRLNHNKYNPVS 832

Db 121 LLEDTPROCRICGGLAMVECRECYDDPDISAGIKQFCKTQVHLHPKRLNHNKYNPVS 180

Qy 833 LPKLDLPDWDWRHGCIPCONMELFAVLCI 860

Db 181 LPKLDLPDWDWRHGCIPCONMELFAVLCI 208

RESULT 6

US-10-264-237-1610

Sequence 1610, Application US/10264237

Publication No.: US20040009491A1

GENERAL INFORMATION:

APPLICANT: Birse et al.

TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

FILE REFERENCE: PA131P1

CURRENT APPLICATION NUMBER: US/10/264,237

CURRENT FILING DATE: 2002-10-04

PRIOR APPLICATION NUMBER: PCT/US01/16450

PRIOR FILING DATE: 2001-05-18

PRIOR APPLICATION NUMBER: US 60/205,515

PRIOR FILING DATE: 2000-05-19

NUMBER OF SEQ ID NOS: 2876

SOFTWARE: Patent in Ver. 3.1

SEQ ID NO 1610

LENGTH: 238

TYPE: PRT

ORGANISM: Homo sapiens

FEATURE: ; NAME/KEY: MISC FEATURE ; LOCATION: (38) ; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids

FEATURE: ; NAME/KEY: MISC FEATURE ; LOCATION: (187) ; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids

FEATURE: ; NAME/KEY: MISC FEATURE ; LOCATION: (227) ; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids

US-10-264-237-1610

Query Match 15.6%; Score 148; DB 15; Length 238; Best Local Similarity 100.0%; Pred. No. 2.9e-137; Matches 148; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 457 PGNSHGLEVGSLAEVKENPPFVGIVRWIQPPGLNEVLAGLEDEACAGTGTGTRGRY 516

Db 39 PGNSHGLEVGSLAEVKENPPFVGIVRWIQPPGLNEVLAGLEDEACAGTGTGTRGRY 98

Qy 517 FTCAKKAFLVKLKS CRDPSRFPASIQVSNQIERCNSLAFLGYSLVVEENTPPKMEKG 576

Db 99 FTCAKKAFLVKLKS CRDPSRFPASIQVSNQIERCNSLAFLGYSLVVEENTPPKMEKG 158

Qy 577 LEIMIGKKGIQGHYNSCYLDTLFLCLF 604

Db 159 LEIMIGKKGIQGHYNSCYLDTLFLCLF 186

RESULT 7

US-09-864-761-34675

Sequence 34675, Application US/09864761

Patent No. US20020048763A1

GENERAL INFORMATION:

APPLICANT: Penn, Sharon G.

APPLICANT: Rank, David R.

APPLICANT: Hanzel, David K.

APPLICANT: Chen, Wensheng

TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR

FILE REFERENCE: Aecmiga-X-1

CURRENT APPLICATION NUMBER: US/09/864,761

CURRENT FILING DATE: 2001-05-23

PRIOR APPLICATION NUMBER: US 60/180,312

PRIOR FILING DATE: 2000-02-04

PRIOR APPLICATION NUMBER: US 60/207,456

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: US 09/632,366

PRIOR FILING DATE: 2000-08-03

PRIOR APPLICATION NUMBER: GB 24263.6

PRIOR FILING DATE: 2000-10-04

PRIOR APPLICATION NUMBER: US 60/236,359

PRIOR FILING DATE: 2000-09-27

PRIOR APPLICATION NUMBER: PCT/US01/00666

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00667

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00664

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00669

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00665

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00668

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00663

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00662

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00661

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00670

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: US 60/234,687

PRIOR FILING DATE: 2000-09-21

PRIOR APPLICATION NUMBER: US 09/608,408

PRIOR FILING DATE: 2000-06-30

PRIOR APPLICATION NUMBER: US 09/774,203

PRIOR FILING DATE: 2001-01-29

NUMBER OF SEQ ID NOS: 49117

SOFTWARE: Anomax Sequence Listing Engine vers. 1.1

SEQ ID NO 34675

LENGTH: 113

TYPE: PRT

ORGANISM: Homo sapiens

FEATURE: ; OTHER INFORMATION: MAP TO AC007728.1 ; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 0.98 ; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.92 ; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.5 ; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.3 ; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1 ; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 0.89 ; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1 ; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.3 ; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.4 ; OTHER INFORMATION: EST\_HUMAN HIT: A1130924.1, EVALUATE 5.00e-62

OTHER INFORMATION: SWISSPROT HIT: Q03164, EVALUE 8.90e-01  
US-09-864-761-34675

Query Match 11.9%; Score 113; DB 9; Length 113;  
Best Local Similarity 100.0%; Pred. No. 5.9e-103; Indels 0; Gaps 0;  
Matches 113; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 387 DFRSSPPLQPPVNSLTENRHFSLTKMPNTNGSIGHSPLSAQSVMEELNAP 446  
Db 1 DFRSSPPLQPPVNSLTENRHFSLTKMPNTNGSIGHSPLSAQSVMEELNAP 60  
  
Qy 447 VQESPLAMPNGSHGLVGSIAEVKENPPFYGVIRWIGQPPGLNEVLGLLE 499  
Db 61 VQESPLAMPNGSHGLVGSIAEVKENPPFYGVIRWIGQPPGLNEVLGLLE 113

RESULT 8  
US-10-424-599-186263  
; Sequence 186263, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 2314-248  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 186263  
; LENGTH: 49  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_139208C.1.pap  
US-10-424-599-186263

Query Match 0.8%; Score 8; DB 15; Length 49;  
Best Local Similarity 100.0%; Pred. No. 18;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 223 VELPPLEI 230  
Db 36 VELPPLEI 43

RESULT 9  
US-10-058-053A-218  
; Sequence 218, Application US/10058053A  
; Publication No. US2003017022A1  
; GENERAL INFORMATION:  
; APPLICANT: University of Utah Research Foundation  
; APPLICANT: Cognetix, Inc.  
; APPLICANT: Jones, Robert M.  
; APPLICANT: Garrett, James E.  
; APPLICANT: Watkins, Maren  
; APPLICANT: Olivera, Baldomero M.  
; TITLE OF INVENTION: B-Superfamily Conotoxins  
; FILE REFERENCE: 2314-248  
; CURRENT APPLICATION NUMBER: US/10/058,053A  
; CURRENT FILING DATE: 2002-07-05  
; PRIOR APPLICATION NUMBER: US 60/264323  
; PRIOR FILING DATE: 2001-01-29  
; NUMBER OF SEQ ID NOS: 340  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 218  
; LENGTH: 78  
; TYPE: PRT  
; ORGANISM: Conus terebra  
; FEATURE:  
; LOCATION: (1)..(122)  
; OTHER INFORMATION: unsure at all xaa locations  
US-10-058-053A-218

Query Match 0.8%; Score 8; DB 14; Length 78;  
Best Local Similarity 100.0%; Pred. No. 28;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 916 LHSLSRR 923  
Db 43 LHSLSRR 50  
  
RESULT 10  
US-10-838-226-218  
; Sequence 218, Application US/10838226  
; Publication No. US20040176278A1  
; GENERAL INFORMATION:  
; APPLICANT: University of Utah Research Foundation  
; APPLICANT: Cognetix, Inc.  
; APPLICANT: Jones, Robert M.  
; APPLICANT: Garrett, James E.  
; APPLICANT: Watkins, Maren  
; APPLICANT: Olivera, Baldomero M.  
; TITLE OF INVENTION: B-Superfamily Conotoxins  
; FILE REFERENCE: 2314-241  
; CURRENT APPLICATION NUMBER: US/10/838,226  
; CURRENT FILING DATE: 2004-05-05  
; PRIOR APPLICATION NUMBER: US 10/058,053  
; PRIOR FILING DATE: 2000-01-29  
; PRIOR APPLICATION NUMBER: US 60/264323  
; PRIOR FILING DATE: 2001-01-29  
; NUMBER OF SEQ ID NOS: 340  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 218  
; LENGTH: 78  
; TYPE: PRT  
; ORGANISM: Conus terebra  
US-10-838-226-218

Query Match 0.8%; Score 8; DB 16; Length 78;  
Best Local Similarity 100.0%; Pred. No. 28;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 916 LHSLSRR 923  
Db 43 LHSLSRR 50

RESULT 11  
US-10-437-963-180681  
; Sequence 180681, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 180681  
; LENGTH: 122  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(122)  
; OTHER INFORMATION: unsure at all xaa locations  
US-10-437-963-180681

; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_78029C.1.pep  
US-10-437-963-180681

Query Match 0.8%; Score 8; DB 16; Length 122;  
Best Local Similarity 100.0%; Pred. No. 42;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 450 SPPLAMPP 457  
|||||  
DB 16 SPPLAMPP 23

## RESULT 12

US-10-424-599-170638  
; Sequence 170638, Application US/10424599  
; Publication No. US20040031072A1

## GENERAL INFORMATION:

; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei

; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

; FILE REFERENCE: 38-21(53223)B

; CURRENT APPLICATION NUMBER: US/10/424,599

; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 285684

; SEQ ID NO 170638

; LENGTH: 143

; TYPE: PRT

; ORGANISM: Glycine max

; FEATURE:

; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_125100C.1.pep

US-10-424-599-170638

Query Match 0.8%; Score 8; DB 15; Length 143;  
Best Local Similarity 100.0%; Pred. No. 49;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 VELLEGR 171  
|||||  
DB 99 VELLEGR 106

## RESULT 13

US-10-424-599-159539  
; Sequence 159539, Application US/10424599  
; Publication No. US20040031072A1

## GENERAL INFORMATION:

; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei

; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

; FILE REFERENCE: 38-21(53223)B

; CURRENT APPLICATION NUMBER: US/10/424,599

; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 285684

; SEQ ID NO 159539

; LENGTH: 188

; TYPE: PRT

; ORGANISM: Glycine max

; FEATURE:

; NAME/KEY: unsure

; LOCATION: (1)..(188)

; OTHER INFORMATION: unsure at all Xaa locations

; FEATURE:

; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_115083C.1.pep

US-10-424-599-159539

Query Match 0.8%; Score 8; DB 15; Length 188;  
Best Local Similarity 100.0%; Pred. No. 63;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 547 QIERCNSL 554  
|||||  
DB 120 QIERCNSL 127

## RESULT 14

US-10-437-963-188186  
; Sequence 188186, Application US/10437963  
; Publication No. US20040123343A1

## GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei

; APPLICANT: Wu, Wei

; APPLICANT: Boukharov, Andrey A.

; APPLICANT: Barbazuk, Brad

; APPLICANT: Li, Ping

; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

; FILE REFERENCE: 38-21(53221)B

; CURRENT APPLICATION NUMBER: US/10/437,963

; CURRENT FILING DATE: 2003-05-14

; NUMBER OF SEQ ID NOS: 204966

; SEQ ID NO 188186

; LENGTH: 371

; TYPE: PRT

; ORGANISM: Oryza sativa

; FEATURE:

; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_84815C.1.pep

US-10-437-963-188186

Query Match 0.8%; Score 8; DB 16; Length 371;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 391 SSPPLQPP 398  
|||||  
DB 221 SSPPLQPP 228

## RESULT 15

US-10-282-122A-60805  
; Sequence 60805, Application US/10282122A  
; Publication No. US20040029129A1

## GENERAL INFORMATION:

; APPLICANT: Wang, Liangsu

; APPLICANT: Zamudio, Carlos

; APPLICANT: Malone, Cheryl

; APPLICANT: Haselbeck, Robert

; APPLICANT: Ohlsen, Kari

; APPLICANT: Zyskind, Judith

; APPLICANT: Wall, Daniel

; APPLICANT: Trawick, John

; APPLICANT: Carr, Grant

; APPLICANT: Yamamoto, Robert

; APPLICANT: Forsyth, R.

; APPLICANT: Xu, H.

; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

; FILE REFERENCE: ELITRA.034A

; CURRENT APPLICATION NUMBER: US/10/282,122A

; CURRENT FILING DATE: 2003-02-20

; PRIOR APPLICATION NUMBER: 60/191,078

; PRIOR FILING DATE: 2000-03-21

; PRIOR APPLICATION NUMBER: 60/206,848

; PRIOR FILING DATE: 2000-05-23

; PRIOR APPLICATION NUMBER: 60/207,727

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: 60/230,335

; PRIOR FILING DATE: 2000-09-06

; PRIOR APPLICATION NUMBER: 60/230,347

; PRIOR FILING DATE: 2000-09-09  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/267,636  
; PRIOR FILING DATE: 2001-02-09  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 78614  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 60805  
; LENGTH: 450  
; TYPE: PRT  
; ORGANISM: Listeria monocytogenes  
US-10-282-122A-60805

Query Match 0.8%; Score 8; DB 15; Length 450;  
Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 45 SIGQYIQD 52  
|||  
Db 312 SIGQYIQD 319

Search completed: April 18, 2005, 09:58:40  
Job time : 59 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 18, 2005, 09:18:24 ; Search time 57 Seconds  
(without alignments)  
1242.841 Million cell updates/sec

Title: US-09-671-687A-3  
Perfect score: 949  
Sequence: 1 MSSGLMSQEKVTSPLYWERRI.....RLLCDAYNCVQSPMTSLYK 949

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 513545 seqs, 74649064 residues

Word size : 0  
Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued Patents\_AA:\*  
1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep:\*  
2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep:\*  
3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep:\*  
4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep:\*  
5: /cgn2\_6/ptodata/1/iaa/PCTUS\_COMB.pep:\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	8	0.8	228	4	US-09-252-991A-22217
2	8	0.8	344	4	US-09-270-767-42367
3	8	0.8	708	4	US-09-252-991A-20021
4	8	0.8	2216	4	US-09-902-540-12221
5	7	0.7	28	3	US-09-227-357-607
6	7	0.7	56	4	US-09-513-999C-6825
7	7	0.7	62	1	US-08-616-732A-24
8	7	0.7	62	3	US-09-037-742B-24
9	7	0.7	108	4	US-09-513-999C-6103
10	7	0.7	112	4	US-09-902-540-14754
11	7	0.7	116	2	US-08-879-995A-4
12	7	0.7	116	3	US-09-215-096-4
13	7	0.7	125	4	US-09-583-110-4643
14	7	0.7	127	4	US-09-513-999C-4224
15	7	0.7	134	4	US-09-107-433-2866
16	7	0.7	135	4	US-09-270-767-36133
17	7	0.7	135	4	US-09-270-767-51350
18	7	0.7	140	4	US-09-270-767-32498
19	7	0.7	140	4	US-09-270-767-47715
20	7	0.7	142	4	US-09-602-777A-334
21	7	0.7	144	4	US-09-248-796A-18078
22	7	0.7	147	1	US-08-688-609-8
23	7	0.7	147	3	US-09-002-832-8
24	7	0.7	151	4	US-09-270-767-44667
25	7	0.7	151	4	US-09-902-540-15459
26	7	0.7	152	4	US-09-621-976-4852
27	7	0.7	153	4	US-09-270-767-56889

28 7 0.7 156 4 US-09-710-279-1222 Sequence 1222, Ap  
29 7 0.7 162 4 US-09-270-767-33034 Sequence 33034, A  
30 7 0.7 162 4 US-09-270-767-36622 Sequence 36622, A  
31 7 0.7 162 4 US-09-270-767-48251 Sequence 48251, A  
32 7 0.7 162 4 US-09-270-767-51839 Sequence 51839, A  
33 7 0.7 163 2 US-08-783-395-5 Sequence 5, Appli  
34 7 0.7 163 3 US-08-600-430-2 Sequence 2, Appli  
35 7 0.7 163 3 US-09-134-001C-4131 Sequence 4131, Ap  
36 7 0.7 163 4 US-09-461-912A-44 Sequence 44, Appl  
37 7 0.7 163 4 US-09-949-016-6077 Sequence 6077, Ap  
38 7 0.7 169 4 US-09-711-164-381 Sequence 381, App  
39 7 0.7 172 4 US-09-270-767-41549 Sequence 41649, A  
40 7 0.7 174 4 US-09-949-016-11208 Sequence 11208, A  
41 7 0.7 181 4 US-09-513-999C-5797 Sequence 5797, Ap  
42 7 0.7 185 4 US-09-583-110-4438 Sequence 4438, Ap  
43 7 0.7 190 4 US-09-328-352-6627 Sequence 6627, Ap  
44 7 0.7 192 4 US-09-107-433-3944 Sequence 3944, Ap  
45 7 0.7 193 4 US-09-252-991A-17828 Sequence 17828, A

## ALIGNMENTS

RESULT 1  
US-09-252-991A-22217  
; Sequence 22217, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 22217  
; LENGTH: 228  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-22217

Query Match 0.8%; Score 8; DB 4; Length 228;  
Best Local Similarity 100.0%; Pred. No. 27;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 VELLEGR 171  
DB 151 VELLEGR 158

RESULT 2  
US-09-270-767-42367  
; Sequence 42367, Application US/09270767  
; Patent No. 6703491  
; GENERAL INFORMATION:  
; APPLICANT: Homburger et al.  
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster  
; FILE REFERENCE: File Reference: 7326-094  
; CURRENT APPLICATION NUMBER: US/09/270,767  
; CURRENT FILING DATE: 1999-03-17  
; NUMBER OF SEQ ID NOS: 62517  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 42367  
; LENGTH: 344  
; TYPE: PRT  
; ORGANISM: Drosophila melanogaster  
US-09-270-767-42367

Query Match 0.8%; Score 8; DB 4; Length 344;

Best Local Similarity 100.0%; Pred. No. 39;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 225 LPPLEINS 232  
| | | | |  
Db 164 LPPLEINS 171

## RESULT 3

US-09-252-991A-20021  
; Sequence 20021, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 20021  
; LENGTH: 708  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-20021

## Query Match

Best Local Similarity 0.8%; Score 8; DB 4; Length 708;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 490 LNEVLGL 497  
| | | | |  
Db 266 LNEVLGL 273

## RESULT 4

US-09-902-540-12221  
; Sequence 12221, Application US/09902540  
; Patent No. 6833447  
; GENERAL INFORMATION:  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Wiegand, Roger C.  
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof  
; FILE REFERENCE: 38-10(15849)B  
; CURRENT APPLICATION NUMBER: US/09/902,540  
; CURRENT FILING DATE: 2001-07-10  
; PRIOR APPLICATION NUMBER: 60/217,883  
; PRIOR FILING DATE: 2000-07-10  
; NUMBER OF SEQ ID NOS: 16825  
; SEQ ID NO 12221  
; LENGTH: 2216  
; TYPE: PRT  
; ORGANISM: Myxococcus xanthus  
US-09-902-540-12221

## Query Match

Best Local Similarity 0.8%; Score 8; DB 4; Length 2216;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 169 EGRGQFT 176  
| | | | |  
Db 395 EGRGQFT 402

## RESULT 5

US-09-227-357-607  
; Sequence 607, Application US/09227357

; Patent No. 6342581  
; GENERAL INFORMATION:  
; APPLICANT: Fischer et al.  
; TITLE OF INVENTION: 123 Human Secreted Proteins  
; FILE REFERENCE: PZ010P1  
; CURRENT APPLICATION NUMBER: US/09/227,357  
; CURRENT FILING DATE: 1999-01-08  
; EARLIER APPLICATION NUMBER: PCT/US98/13684  
; EARLIER FILING DATE: 1998-07-07  
; EARLIER APPLICATION NUMBER: 60/051,926  
; EARLIER FILING DATE: 1997-07-08  
; EARLIER APPLICATION NUMBER: 60/052,793  
; EARLIER FILING DATE: 1997-07-08  
; EARLIER APPLICATION NUMBER: 60/051,925  
; EARLIER FILING DATE: 1997-07-08  
; EARLIER APPLICATION NUMBER: 60/051,929  
; EARLIER FILING DATE: 1997-07-08  
; EARLIER APPLICATION NUMBER: 60/052,803  
; EARLIER FILING DATE: 1997-07-08  
; EARLIER APPLICATION NUMBER: 60/052,732  
; EARLIER FILING DATE: 1997-07-08  
; EARLIER APPLICATION NUMBER: 60/051,931  
; EARLIER FILING DATE: 1997-07-08  
; EARLIER APPLICATION NUMBER: 60/051,932  
; EARLIER FILING DATE: 1997-07-08  
; EARLIER APPLICATION NUMBER: 60/051,916  
; EARLIER FILING DATE: 1997-07-08  
; EARLIER APPLICATION NUMBER: 60/051,930  
; EARLIER FILING DATE: 1997-07-08  
; EARLIER APPLICATION NUMBER: 60/051,918  
; EARLIER FILING DATE: 1997-07-08  
; EARLIER APPLICATION NUMBER: 60/051,920  
; EARLIER FILING DATE: 1997-07-08  
; EARLIER APPLICATION NUMBER: 60/052,733  
; EARLIER FILING DATE: 1997-07-08  
; EARLIER APPLICATION NUMBER: 60/052,795  
; EARLIER FILING DATE: 1997-07-08  
; EARLIER APPLICATION NUMBER: 60/051,919  
; EARLIER FILING DATE: 1997-07-08  
; EARLIER APPLICATION NUMBER: 60/051,928  
; EARLIER FILING DATE: 1997-07-08  
; EARLIER APPLICATION NUMBER: 60/055,722  
; EARLIER FILING DATE: 1997-08-18  
; EARLIER APPLICATION NUMBER: 60/055,723  
; EARLIER FILING DATE: 1997-08-18  
; EARLIER APPLICATION NUMBER: 60/055,948  
; EARLIER FILING DATE: 1997-08-18  
; EARLIER APPLICATION NUMBER: 60/055,949  
; EARLIER FILING DATE: 1997-08-18  
; EARLIER APPLICATION NUMBER: 60/055,953  
; EARLIER FILING DATE: 1997-08-18  
; EARLIER APPLICATION NUMBER: 60/055,950  
; EARLIER FILING DATE: 1997-08-18  
; EARLIER APPLICATION NUMBER: 60/055,947  
; EARLIER FILING DATE: 1997-08-18  
; EARLIER APPLICATION NUMBER: 60/055,964  
; EARLIER FILING DATE: 1997-08-18  
; EARLIER APPLICATION NUMBER: 60/056,360  
; EARLIER FILING DATE: 1997-08-18  
; EARLIER APPLICATION NUMBER: 60/055,684  
; EARLIER FILING DATE: 1997-08-18  
; EARLIER APPLICATION NUMBER: 60/055,984  
; EARLIER FILING DATE: 1997-08-18  
; EARLIER APPLICATION NUMBER: 60/055,954  
; EARLIER FILING DATE: 1997-08-18  
; EARLIER APPLICATION NUMBER: 60/058,785  
; EARLIER FILING DATE: 1997-09-12  
; EARLIER APPLICATION NUMBER: 60/058,664  
; EARLIER FILING DATE: 1997-09-12  
; EARLIER APPLICATION NUMBER: 60/058,660  
; EARLIER FILING DATE: 1997-09-12  
; EARLIER APPLICATION NUMBER: 60/058,661  
; EARLIER FILING DATE: 1997-09-12

NUMBER OF SEQ ID NOS: 672  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 607  
LENGTH: 28  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-227-357-607

Query Match 0.7%; Score 7; DB 3; Length 28;  
Best Local Similarity 100.0%; Pred. No. 38;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 732 INSNLKF 738  
Db 5 INSNLKF 11  
|||||

## RESULT 6

US-09-513-999C-6825  
Sequence 6825, Application US/09513999C  
Patent No. 6783961  
GENERAL INFORMATION:  
APPLICANT: Dumas Milne Edwards, J.B.  
APPLICANT: Duclert, A.  
APPLICANT: Giordano, J.Y.  
TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.  
Patent No. 6783961  
FILE REFERENCE: 59.US2.REG  
CURRENT APPLICATION NUMBER: US/09/513.999C  
CURRENT FILING DATE: 2000-02-24  
PRIOR APPLICATION NUMBER: US 60/122,487  
PRIOR FILING DATE: 1999-02-26  
NUMBER OF SEQ ID NOS: 36881  
SOFTWARE: Patent.Pm  
SEQ ID NO 6825  
LENGTH: 56  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-513-999C-6825

QY 321 GDKGSSS 327  
Db 20 GDKGSSS 26  
|||||

## RESULT 7

US-08-616-732A-24  
Sequence 24, Application US/08616732A  
Patent No. 5770690  
GENERAL INFORMATION:  
APPLICANT: Bitler, Catherine Mastroni  
APPLICANT: Bowersox, Stephen Scott  
APPLICANT: Crea, Roberto  
APPLICANT: Demo, Susan Dunham  
APPLICANT: Horne, William A.  
APPLICANT: Zhou, Mei  
TITLE OF INVENTION: Bax Omega Protein and Methods  
NUMBER OF SEQUENCES: 27  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dehlinger & Associates  
STREET: 350 Cambridge Avenue, Suite 250  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94306  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/616,732A  
FILING DATE: 15-MAR-1996  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/495,042  
FILING DATE: 27-JUN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Sholtz, Charles K.  
REGISTRATION NUMBER: 38,615  
REFERENCE/DOCKET NUMBER: 5865-0017.30  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 324-0880  
TELEFAX: (415) 324-0960  
INFORMATION FOR SEQ ID NO: 24:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 62 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: translation of third open reading  
frame in Fig. 2  
INDIVIDUAL ISOLATE: frame in Fig. 2  
US-08-616-732A-24

Query Match 0.7%; Score 7; DB 1; Length 62;  
Best Local Similarity 100.0%; Pred. No. 79;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 739 AEAPSCL 745  
Db 49 AEAPSCL 55  
|||||

## RESULT 8

US-09-037-742B-24  
Sequence 24, Application US/09037742B  
Patent No. 6140484  
GENERAL INFORMATION:  
APPLICANT: Bitler, Catherine Mastroni  
APPLICANT: Bowersox, Stephen Scott  
APPLICANT: Crea, Roberto  
APPLICANT: Demo, Susan Dunham  
APPLICANT: Horne, William A.  
APPLICANT: Zhou, Mei  
TITLE OF INVENTION: Bax Omega Protein and Methods  
NUMBER OF SEQUENCES: 27  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dehlinger & Associates  
STREET: 350 Cambridge Avenue, Suite 250  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94306  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/037,742B  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/616,732  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Sholtz, Charles K.  
REGISTRATION NUMBER: 38,615

REFERENCE/DOCKET NUMBER: 5865-0017.30  
TELEPHONE: (415) 324-0880  
TELEFAX: (415) 324-0960  
INFORMATION FOR SEQ ID NO: 24:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 62 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE: translation of third open reading  
INDIVIDUAL ISOLATE: frame in Fig. 2  
INDIVIDUAL ISOLATE: frame in Fig. 2  
US-09-037-742B-24

Query Match 0.7%; Score 7; DB 3; Length 62;  
Best Local Similarity 100.0%; Pred. No. 79;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 739 AEAPSL 745  
Db 49 AEAPSL 55

## RESULT 9

US-09-513-999C-6103  
Sequence 6103, Application US/09513999C  
Patent No. 6783961  
GENERAL INFORMATION:  
APPLICANT: Dumas Milne Edwards, J.B.  
APPLICANT: Duclert, A.Y.  
TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.  
FILE REFERENCE: 59, US2.REG  
CURRENT APPLICATION NUMBER: US/09/513,999C  
CURRENT FILING DATE: 2000-02-24  
PRIOR APPLICATION NUMBER: US 60/122,487  
PRIOR FILING DATE: 1999-02-26  
NUMBER OF SEQ ID NOS: 36681  
SOFTWARE: Patent.pm  
SEQ ID NO 6103  
LENGTH: 108  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-513-999C-6103

Query Match 0.7%; Score 7; DB 4; Length 108;  
Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 393 PPLQPPP 399  
Db 100 PPLQPPP 106

## RESULT 10

US-09-540-14754  
Sequence 14754, Application US/09902540  
Patent No. 6833447  
GENERAL INFORMATION:  
APPLICANT: Goldman, Barry S.  
APPLICANT: Hinkle, Gregory J.  
APPLICANT: Slater, Steven C.  
APPLICANT: Wiegand, Roger C.  
TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof  
FILE REFERENCE: 38-10(115849)B  
CURRENT APPLICATION NUMBER: US/09/902,540  
CURRENT FILING DATE: 2001-07-10  
PRIOR APPLICATION NUMBER: 60/217,883

PRIOR FILING DATE: 2000-07-10  
NUMBER OF SEQ ID NOS: 16825  
SEQ ID NO 14754  
LENGTH: 112  
TYPE: PRT  
ORGANISM: Myxococcus xanthus  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (1)..(112)  
OTHER INFORMATION: unsure at all Xaa locations  
US-09-902-540-14754

Query Match 0.7%; Score 7; DB 4; Length 112;  
Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 696 RSAGQKV 702  
Db 73 RSAGQKV 79

## RESULT 11

US-08-879-995A-4  
Sequence 4, Application US/08879995A  
Patent No. 5985606  
GENERAL INFORMATION:  
APPLICANT: Hillman, Jennifer L.  
APPLICANT: Lal, Preeti  
APPLICANT: Kaser, Matthew R.  
TITLE OF INVENTION: HUMAN PREPROTACHYKININ B  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/879,995A  
FILING DATE: Herewith  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0326 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-855-0555  
TELEFAX: 415-845-4166  
TELEX:  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 116 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: GenBank  
CLONE: 205725  
US-08-879-995A-4

Query Match 0.7%; Score 7; DB 2; Length 116;  
Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 563 VVENTP 569  
|||||||  
Db 104 VVENTP 110

## RESULT 12

US-09-215-096-4  
; Sequence 4, Application US/09215096  
; Patent No. 6008194  
; GENERAL INFORMATION:  
; APPLICANT: Hillman, Jennifer L.  
; APPLICANT: Lal, Preeti  
; APPLICANT: Kaser, Matthew R.  
; TITLE OF INVENTION: HUMAN PREPROTACHYKININ B  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Drive  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Fast-Seq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/215,096  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/879,995  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Billings, Lucy J.  
; REGISTRATION NUMBER: 36,749  
; REFERENCE/DOCKET NUMBER: PF-0326 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-855-0555  
; TELEFAX: 415-845-4166  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 116 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; LIBRARY: GenBank  
; CLONE: 205725  
; US-09-215-096-4

Query Match 0.7%; Score 7; DB 3; Length 116;  
Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 563 VVENTP 569  
|||||||  
Db 104 VVENTP 110

## RESULT 13

US-09-583-110-4643  
; Sequence 4643, Application US/09583110  
; Patent No. 6699703  
; GENERAL INFORMATION:  
; APPLICANT: Lynn Doucette-Stamm et al.  
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus  
; FILE REFERENCE: PATH00-07A  
; CURRENT APPLICATION NUMBER: US/09/583,110  
; CURRENT FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: US 09/107,433  
; PRIOR FILING DATE: 1998-06-30  
; PRIOR APPLICATION NUMBER: US 60/085,131  
; PRIOR FILING DATE: 1998-05-12  
; PRIOR APPLICATION NUMBER: US 60/051,553  
; PRIOR FILING DATE: 1997-07-02  
; NUMBER OF SEQ ID NOS: 5322  
; SEQ ID NO 4643  
; LENGTH: 125  
; TYPE: PRT  
; ORGANISM: Streptococcus pneumoniae  
; US-09-583-110-4643

Query Match 0.7%; Score 7; DB 4; Length 125;  
Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 688 RVEPLK 694  
|||||||  
Db 112 RVEPLK 118

## RESULT 14

US-09-513-999C-4224  
; Sequence 4224, Application US/09513999C  
; Patent No. 6783961  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Duclert, A.  
; APPLICANT: Giordano, J.Y.  
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.  
; Patent No. 6783961  
; FILE REFERENCE: 59.US2.REG  
; CURRENT APPLICATION NUMBER: US/09/513,999C  
; CURRENT FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/122,487  
; PRIOR FILING DATE: 1999-02-26  
; NUMBER OF SEQ ID NOS: 36681  
; SOFTWARE: Patent.pm  
; SEQ ID NO 4224  
; LENGTH: 127  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SIGNAL  
; LOCATION: -23...-1  
; OTHER INFORMATION: score 4.1  
; OTHER INFORMATION: seq LHLILLLVFVAT/LD  
; FEATURE:  
; NAME/KEY: UNSURE  
; LOCATION: 5  
; OTHER INFORMATION: Xaa=Arg or Trp  
; FEATURE:  
; NAME/KEY: UNSURE  
; LOCATION: 91  
; OTHER INFORMATION: Xaa=any one of the twenty amino acids  
; US-09-513-999C-4224

Query Match 0.7%; Score 7; DB 4; Length 127;  
Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 253 LFGKESL 259  
|||||||  
Db 31 LFGKESL 37

## RESULT 15

US-09-107-433-2866  
; Sequence 2866, Application US/09107433  
; Patent No. 6800744  
; GENERAL INFORMATION:  
; APPLICANT: Lynn A Doucette-Stamm and David Bush

;; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID  
;; SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE FOR DIAGN  
;; THERAPEUTICS  
;;  
;; NUMBER OF SEQUENCES: 5206  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: GENOME THERAPEUTICS CORPORATION  
;; STREET: 100 Beaver Street  
;; CITY: Waltham  
;; STATE: Massachusetts  
;; COUNTRY: USA  
;; ZIP: 02354  
;;  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: CD-ROM ISO9660  
;; COMPUTER: <Unknown>  
;; OPERATING SYSTEM: <Unknown>  
;; SOFTWARE: <Unknown>  
;;  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/09/107,433  
;; FILING DATE: 30-Jun-1998  
;;  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: 60/ 085131  
;; FILING DATE: May 12, 1998  
;; APPLICATION NUMBER: 60/051553  
;; FILING DATE: July 2, 1997  
;;  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Aziniello, Pamela Deneke  
;; REGISTRATION NUMBER: 40,489  
;; REFERENCE/DOCKET NUMBER: GTC-011  
;;  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (781)893-5007  
;; TELEFAX: (781)893-8277  
;;  
;; INFORMATION FOR SEQ ID NO: 2866:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 134 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
;; HYPOTHETICAL: YES  
;; ORIGINAL SOURCE:  
;; ORGANISM: Streptococcus pneumoniae  
;; FEATURE:  
;; NAME/KEY: misc feature  
;; LOCATION: (8) LOCATION 1...134  
;; SEQUENCE DESCRIPTION: SEQ ID NO: 2866:  
US-09-107-433-2866

Query Match 0.7%; Score 7; DB 4; Length 134;  
Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 688 RVEPLK 694  
Db 121 RVEPLK 127

Search completed: April 18, 2005, 09:42:12  
Job time : 58 secs